The Rare Decays $B^0_{(s)} \to \mu^+\mu^-$ and Z Boson Production at LHCb and Radiation Damage in its Silicon Tracker

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Christian Elsasser

von

Zürich ZH
Aarau AG
Unterkulm AG

Promotionskomitee

Prof. Dr. Ulrich Straumann (Vorsitz)
Prof. Dr. Laura Baudis
Dr. Katharina Müller
Dr. Olaf Steinkamp

Zürich, 2014

Contact:
Christian Elsasser
elsasser@cern.ch

Abstract

The LHCb experiment at the Large Hadron Collider is dedicated to the study of physics involving beauty and charm quarks. Its detector – designed as a single-arm forward spectrometer – uses tracking stations upstream and downstream of a dipole magnet to measure the trajectories and the momenta of charged particles. Parts of this tracking system are built from silicon microstrip detectors. This thesis describes the monitoring of the radiation damage in these detectors using leakage current measurements and so-called Charge Collection Efficiency scans performed in regular intervals during data taking. The results of this monitoring during the first LHC data taking period from 2010 to 2013 are presented and predictions for the future evolution of the detector ageing are discussed.

A main goal of LHCb is the measurement of rare decays. As these decays are heavily suppressed in the Standard Model, their properties, for example the branching fraction, angular distributions or CP asymmetries, can be significantly modified by contributions from physics beyond the Standard Model. "Golden" channels among such indirect searches are the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$. This thesis describes for these two decays the determination of the expected signal response in terms of the invariant dimuon mass and a multivariate classifier based on topological and kinematical decay properties. This calibration is based on known dimuon resonances such as $J/\psi \to \mu^+\mu^-$, $\Upsilon(1S) \to \mu^+\mu^-$ or $Z \to \mu^+\mu^-$ and charmless two-body decays of B_s^0 and B^0 mesons into pairs of charged kaons and pions. The developed techniques were also applied to the determination of the signal yield of normalisation channels, the estimation of background contamination and the comparison of selection variables between data and simulation. These investigations were a part of the analysis that led to the first evidence of the decay $B_s^0 \to \mu^+\mu^-$ and the measurement of its branching fraction.

Besides measurements involving beauty and charm quarks, LHCb is well suited to measure particles produced in hadron collisions at small angles with respect to the beam axis. An interesting measurement to investigate the dynamics of partons inside nuclei is the study of electroweak gauge boson production in proton-lead collisions. A first measurement of Z boson production based on proton-lead data collected by LHCb in 2013 is also presented in this thesis. It led to the first observation ever of Z boson production in proton-nuclei collisions.

Zusammenfassung

Das LHCb-Experiment am Large Hadron Collider hat zum Ziel, Eigenschaften von Teilchen, welche Charm- oder Beauty-Quarks enthalten, zu studieren. Der dabei verwendete Detektor ist als einarmiges Spektrometer entlang der Strahlachse des LHC aufgebaut. Er verfügt über einen Dipolmagneten mit vor- und nachgelagerten Spurdetektoren zur Bestimmung der Impulse und Spuren von geladenen Teilchen. Teile dieser Spurdetektoren sind als Siliziumstreifendetektoren gebaut. Die vorliegende Arbeit beschreibt die Überwachung von Strahlungschäden in diesen Detektoren mittels regelmässiger Messungen der Leckströme im Silizium sowie der benötigten Spannung, welche an den Siliziumsensoren angelegt werden muss, um aus diesen thermisch erzeugte Ladungsträger zu entfernen. In der vorliegenden Arbeit werden die Resultate dieser Überwachungsmessungen während der ersten Messperiode am LHC zwischen 2010 und 2013 präsentiert sowie darauf basierende Vorhersagen der künftigen Strahlungsschäden erläutert.

Ein Hauptziel des LHCb-Experiments ist die Messung von sogenannten Seltenen Zerfällen. Diese Zerfälle sind im Standard-Modell stark unterdrückt. Deshalb können sich ihre Eigenschaften, wie zum Beispiel das Verzweigungsverhältnis, die Winkelverteilung der Zerfallsteilchen sowie die Grösse der CP-Asymmetrie, unter dem Einfluss von Dynamiken, welche nicht durch das Standard-Modell beschrieben werden, signifikant verändern. Zwei "goldene" Zerfallskanäle sind die Zerfälle $B_s^0 \to \mu^+\mu^-$ und $B^0 \to \mu^+\mu^-$. Diese Arbeit beschreibt für die zwei Zerfälle die Bestimmung der Signalverteilung der invarianten Masse der beiden Myonen im Zerfall sowie einer multivariaten Variablen, welche auf topologischen und kinematischen Eigenheiten der Zerfälle basiert. Diese Variable hat das Ziel Signalzerfälle von Hintergrundprozessen zu unterscheiden. Die Kalibrationen beruhen auf bekannten Resonanzen, welche in zwei Myonen zerfallen, wie zum Beispiel $J/\psi \to \mu^+\mu^-$, $\Upsilon(1S) \to \mu^+\mu^-$ oder $Z \to \mu^+\mu^-$ sowie Zwei-Körper-Zerfällen von B_s^0- und B^0-Mesonen in zwei geladene Pionen oder Kaonen. Die entwickelten Algorithmen wurden auch zur Bestimmung der Anzahl Signalereignisse in Normalisationskanälen, zur Abschätzung der Anzahl Hintergrundereignisse sowie zum Vergleich der Verteilungen von zur Signalselektion geeigneten Grössen zwischen Simulationen und echten Daten aus Proton-Proton-Kollisionen verwendet. Diese Untersuchungen bildeten einen Hauptteil der Analyse, welche zum ersten Beweis für den Zerfall $B_s^0 \to \mu^+\mu^-$ und der Bestimmung dessen Verzweigungsverhältnisses geführt hat.

Neben Messungen mit Teilchen, welche Beauty- oder Charm-Quarks enthalten, kann LHCb die Produktion von Teilchen, welche in Kollisionen von Hadronen unter kleinen Winkeln zur Strahlachse des LHC entstehen, messen. Ein Beispiel dafür ist die Untersuchung der Dynamik von Partonen in einem Atomkern durch die Messung des Wirkungsquerschnitts von Eichbosonen der Elektroschwachen Wechselwirkung in Proton-Blei-Kollisionen. Eine erste Messung des Wirkungsquerschnitts des Z-Bosons, welche auf Daten aus Proton-Blei-Kollisionen in 2013 basiert, wird

ebenfalls in dieser Arbeit präsentiert. Diese Messung ist gleichzeitig auch die erste Beobachtung von Z-Bosonen, welche in Kollisionen von Protonen mit schwereren Atomkernen produziert worden sind.

Acknowledgement

Gratitude where gratitude is due!

Contents

Introduction

In his address to the conference of the British Association for the Advancement of Science in 1900, Lord Kelvin announced: "There is nothing new to be discovered in physics now. All that remains, is more and more precise measurement."[1] He made this statement at a time when the experiments suggested for the first time a substructure of the atom, *i.e.* the discovery of the electron by Thomson in 1897, the discovery of α particles by Rutherford and Soddy in 1903 and finally the discovery of the atom nucleus by Rutherford in 1909. The experiment by Thomson marked the start of series of discoveries of the elementary particles known today, with the latest addition being the discovery of Higgs boson in 2012. This latest discovery marked the highlight of the first running period of the Large Hadron Collider (LHC) and its experiments from 2010 to 2013. Besides direct searches for new particles, a major goal of the LHC is the precise measurement of a wide range of observables allowing to search for tensions with predictions from the fundamental theory of elementary particle physics, known as the Standard Model. Despite Kelvin's claim, a more and more precise measurement can ultimately lead to a discovery.
A dedicated experiment at the LHC that performs these kinds of indirect searches is the LHCb experiment. This experiment is introduced in Chapter 3 of this thesis after a short review of the basic concepts of particle physics, with special emphasis on Heavy Flavour physics in Chapter 1 and a short introduction to CERN and the LHC in Chapter 2.
To perform precision measurements, the LHCb experiment must have an excellent performance, which needs to be constantly monitored and understood. The second part of the thesis is devoted to one of these aspects as Chapters 4 and 5 discuss the monitoring of the radiation damage in the Silicon Tracker. This monitoring allows to extrapolate the long-term behaviour of this component of the detector and to ensure a good performance in the future.
The third part of the thesis describes contributions to the search for the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in LHCb. This is an example of an indirect search for physics not described by the Standard Model by precisely measuring the branching fractions of these decays and looking for a possible tension with the Standard Model predictions. Chapter 6 gives a short review of theoretical aspects of the two decays and explains the search strategy in LHCb. Chapter 7 discusses a multivariate classifier based on topological and kinematical properties of the decay. This classifier is used to distinguish the signal decays from background processes. Its calibration is performed with charmless two-body decays of B^0 and B_s^0 mesons in data, which allows to precisely determine the response of the classifier from data to keep systematical uncertainties as small as possible. Chapter 8 describes the calibration of the invariant dimuon mass distribution for signal.

[1] The attribution of this quote to Lord Kelvin is disputed as no primary source is known. A similar statement is also assigned to Albert Michelson.

This calibration is also mostly done using data to avoid possible biases due to discrepancies between the actual detector response and its description in simulation. Chapter 9 concludes the third part by summarising the latest results in the search of $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ and discussing briefly possible interpretations of these results.
The final part of the thesis is devoted to the first observation of Z boson production in proton-nucleus collisions. This measurement can have – due to the geometry of LHCb – a significant impact on a better understanding of the dynamics inside atomic nuclei. Chapter 10 describes the theoretical foundation of the measurement while Chapter 11 presents the measurement of the Z production cross section in proton-lead collisions in LHCb.

I

The Basics of Particle Physics, CERN, the LHC and LHCb

Three quarks for Muster Mark!
Sure he hasn't got much of a bark
And sure any he has it's all beside the mark.
But O, Wreneagle Almighty, wouldn't un be a sky of a lark
To see that old buzzard whooping about for uns shirt in the dark
And he hunting round for uns speckled trousers around by Palmer-stown Park?
Hohohoho, moulty Mark!

Finnegans Wake, Chapter 24, James A. A. Joyce (1882 - 1941)

1

The Standard Model of Particle Physics

The elementary particles, *i.e.* the particles containing no constituents nor having substructure, and their interactions are described by the Standard Model of particle physics (SM). This description of the behaviour of the basic building blocks of nature has so far been very successful in precisely explaining a wide range of phenomena, from low energy processes such as nuclear decays up to phenomena occurring at the very high energy scales of particle colliders.
The Standard Model comprises Quantum Chromodynamics (QCD) [90, 130, 135] describing the strong force, and the electroweak theory [154, 212, 229] – also called "Glashow-Weinberg-Salam" (GWS) model[1], describing the combination of the weak and electromagnetic forces.
Despite its success in describing the phenomena of these three fundamental forces the SM cannot be a complete theory. There are several phenomena that are not explained by the SM such as Dark Matter or the masses of the neutrinos. Also gravity as the fourth fundamental force is not incorporated in this theory.
This chapter gives a brief overview of the SM. More detailed aspects in context of the decays $B^0_{(s)} \to \mu^+\mu^-$ and of the production of Z bosons in proton-lead collisions are covered in Chapters 6 and 10. Details on the SM beyond this introduction can be found for example in Refs [167], [130], [135] and [90].

1.1 The Particle Content of the Standard Model

The mathematical formalism used in the Standard Model is Quantum Field Theory, where all particles are considered as excitations from the vacuum state of the corresponding fields.
The elementary particles in the SM can be grouped according to their spin. Particles with a full integer spin, and therefore obeying the Bose-Einstein statistics [88], are called *bosons* while those

[1] Sheldon Lee Glashow, Steven Weinberg and Abdus Salam received in 1979 the Nobel Prize in Physics "for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current – describing the electromagnetic as well as the weak force".

with half integer spin are described by the Fermi-Dirac statistics [129, 144] and called *fermions*. Furthermore, for each particle there is an anti-particle having the same properties except opposite charge-like quantum numbers. In some cases, a particle (*e.g.* the photon) is its own anti-particle.

1.1.1 Fermions

Fermions, which are mainly responsible for the properties of matter, can be split into two groups: *quarks* and *leptons*.
Both groups consist of three families (or generations) with two members each. These six different types of quarks and leptons are called flavours. The three quark- and lepton-families only differ in the masses of the particles (*cf.* Tab. 1.1), while the two members of a given family have an electric charge difference $|\Delta Q| = 1$ and form in the case of left-handed quarks or leptons (and the right-handed anti-quarks and -leptons) a weak isospin doublet with components $T_3 = \pm 1/2$. The *weak isospin* is a charge of the weak interaction and is only carried by left-handed fermions[2]. Right-handed fermions build singlets with respect to the weak isospin.
Fermions carry further a *weak hypercharge* Y. This is related to the weak isospin component and the electric charge, Q, via a modified version of the Gell-Mann-Nishijima formula $Q = T_3 + Y/2$ [152, 203].

Table 1.1: Summary of the elementary fermions and their masses[a] [81].

Class	**1st family**	**2nd family**	**3rd family**
Quarks	up (u) $2.3\,\mathrm{MeV}/c^2$	charm (c) $1.28\,\mathrm{GeV}/c^2$	top (t) $173.1\,\mathrm{GeV}/c^2$
	down (d) $4.8\,\mathrm{MeV}/c^2$	strange (s) $95\,\mathrm{MeV}/c^2$	bottom/beauty (b) $4.18\,\mathrm{GeV}/c^2$
Leptons	Electron (e) $511.0\,\mathrm{keV}/c^2$	Muon (μ) $105.6\,\mathrm{MeV}/c^2$	Tau (τ) $1.777\,\mathrm{GeV}/c^2$
	e neutrino[b] $< 2\ \mathrm{eV}/c^2$	μ neutrino[b] $< 0.19\,\mathrm{MeV}/c^2$	τ neutrino[b] $< 18.2\,\mathrm{MeV}/c^2$

[a] Due to the confinement of quarks inside hadrons all quark masses except the top quark mass have been not directly measured, but determined indirectly via their influence on hadronic properties. The numbers quoted here have been calculated based on the minimal subtraction ($\overline{\mathrm{MS}}$) scheme [222, 230]. In the case of the three lightest quarks only lattice calculations have been used [81].

[b] All upper limits correspond to the 95 % C.L.

The important difference between quarks and leptons is that the former carry a *colour-charge* while the latter do not. The colour charge belongs to the strong interaction. Its name comes from the fact that there are three different charge types, called red, green and blue, which behave similarly to the spectral colours in nature: a combination of all three charges (red, green and blue) or a charge and its anti-charge (*e.g.* a red-charged quark and an anti-quark carrying an anti-red charge) are considered as colour-less. Only colour-less objects can be observed as free particles in nature, which is also the reason that free quarks are unobservable. This leads to

[2] A fermion is left-handed if its spin projection on its momentum $\vec{s} \cdot \vec{p}/(|\vec{s}| \cdot |\vec{p}|)$ is -1.

important consequences for the spectrum of objects composite of quarks which will be discussed in Sec. 1.3.
Each quark type consists of three colour fields, which form a triplet for the strong interaction, whereas leptons as colour-less objects form singlets of the strong interaction. A summary of the charges carried by fermions is shown in Tab. 1.2.

Table 1.2: Summary of the charges carried by the elementary fermions. T_3 is the weak isospin component, Y the weak hypercharge and Q the electric charge. U, D, E and N correspond to the up- and down-type quarks as well as to the charged leptons and the neutrinos. Right-handed neutrinos N_R do not carry any charge. Hence, if they exist, they will not participate in any interaction described by the SM.

Class		**Colour**	$\boldsymbol{T_3}$	$\boldsymbol{Y}$	$\boldsymbol{Q}$
Quarks	$Q_L = \begin{pmatrix} U \\ D \end{pmatrix}_L$	Yes	$+1/2$ $-1/2$	$+1/3$	$+2/3$ $-1/3$
	U_R		0	$+4/3$	$+2/3$
	D_R		0	$-2/3$	$-1/3$
Leptons	$L_L = \begin{pmatrix} N \\ E \end{pmatrix}_L$	No	$+1/2$ $-1/2$	-1	0 -1
	E_R		0	-2	-1
	N_R		0	0	0

1.1.2 Gauge Bosons

In Quantum Field Theory, interactions are mediated by the exchange of *gauge bosons*. The gauge bosons of the three forces in the SM are all spin-1 particles. Their properties are described in Tab. 1.3.
The electromagnetic force is mediated by the massless photon (γ). As it is electrically neutral, it cannot self-interact. The Z and W bosons of the weak force are massive and as carriers of weak isospin and hypercharge they self-interact with triple or quartic vertices. Also the gluons (g) as gauge bosons of the strong interaction carry colour-charge and are therefore self-interacting allowing triple and quartic vertices. There are eight gluons different in colour charge[3]. A coulorless gluon cannot exist as it would give the strong interaction – in contrast to the observations – an infinite reach.

1.1.3 Higgs Boson

The final elementary particle in the SM is the *Higgs boson*. It is in contrast to the gauge bosons not an excitation of a force field, but of the Higgs field (*cf.* Sec. 1.2.3). The interplay of this field with the elementary particles gives them a mass, which is proportional to their coupling strength to the Higgs field. The Higgs boson is in contrast to the gauge bosons a spin-0 particle.

[3] The strong interaction is governed by the symmetry group $SU(3)$ as described in Sec. 1.2. This group has eight $(N^2 - 1)$ generators and thus there are eight gauge fields.

Table 1.3: Summary of the properties of the elementary bosons in the SM. T_3 is the weak isospin component, Y the weak hypercharge and Q the electric charge. The masses are from Ref. [81] except the mass of the Higgs boson taken from Ref. [16].

Field/Interaction	**Particle**	**Colour**	T_3	Y	Q	m
Electromagnetic	Photon (γ)	No	0	0	0	0
Weak	Z	No	0	0	0	$91.19\,\mathrm{GeV}/c^2$
	$W^{\pm}$	No	±1	0	±1	$80.38\,\mathrm{GeV}/c^2$
Strong	Gluon (g)	Yes	0	0	0	0
Higgs field	Higgs (H)	No	0	0	0	$125.4\ \mathrm{GeV}/c^2$

On July, 4, 2012 the two LHC experiments CMS and ATLAS announced the discovery of a boson with a mass of about $125.4\,\mathrm{GeV}/c^2$ [12, 111]. All properties measured so far are compatible with the actual Higgs boson of the SM[4] [14, 15, 115, 116, 176, 177].

1.2 The Dynamics of the Standard Model

The dynamics of the Standard Model are described by a Lagrange density $\mathcal{L}_{\mathrm{SM}}$. It can be split into separate parts responsible for different aspects of the model

$$\mathcal{L}_{\mathrm{SM}} \quad = \quad \mathcal{L}_{\mathrm{int}} + \mathcal{L}_{\mathrm{gauge}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}. \tag{1.1}$$

It incorporates the different fields of the elementary particles and has an internal symmetry of type $SU(3)_C \times SU(2)_L \times U(1)_Y$.
The $SU(3)_C$ part of this symmetry governs the behaviour of the strong interaction as it is the symmetry according to the Noether theorem [204] associated to the conservation of the colour charges. It describes the invariance of the Lagrange density under $SU(3)$ transformation of the triplet consisting of the three quark colour fields.
The $SU(2)_L \times U(1)_Y$ part describes the nature of the electroweak interactions: the $SU(2)_L$ invariance is associated to the weak isospin and acts only on the isospin doublets, while the $U(1)_Y$ invariance is connected to the weak hypercharge and acts on all fields carrying it.

[4] The 2013 Nobel Prize in Physics has been awarded to Peter Higgs and François Englert "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider".

1.2.1 The Dynamics and Interactions of the Fermions

The dynamics of the fermion fields ψ and the interaction between them and the gauge bosons are described by

$$\begin{aligned}
\mathcal{L}_{\text{int}} &= \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} \\
&= \sum_{f\in\{\text{I,II,III}\}} \sum_{\psi\in\{L_L^f, E_R^f, Q_L^f, U_R^f, D_R^f\}} \overline{\psi}\gamma^\mu \left(i\partial_\mu - \frac{g'}{2} Y_W B_\mu \right) \psi \\
&- \sum_{f\in\{\text{I,II,III}\}} \sum_{\psi\in\{Q_L^f, L_L^f\}} \overline{\psi} \frac{g}{2} \tau^i W_\mu^i \psi \\
&- \sum_{f\in\{\text{I,II,III}\}} \sum_{\psi\in\{Q_L^f, U_R^f, D_R^f\}} \overline{\psi}\gamma^\mu g_s G_\mu^a T^a \psi, \qquad (1.2)
\end{aligned}$$

where γ^μ are the gamma matrices, g', g and g_s the coupling constants of the weak hypercharge, the weak isospin and the colour charge, Y_W the weak hypercharge, B_μ the corresponding gauge field. τ^i are the three generating matrices of the $SU(2)$ symmetry (the Pauli matrices) and W_μ^i the three gauge fields of the weak isospin. Analogously, T^a are the eight generating matrices of the $SU(3)$ symmetry and G_μ^a are the eight gluon fields. Finally, {I,II,III} are the indices of the three fermion families.

1.2.2 The Dynamics and Self-Interactions of the Gauge Bosons

The dynamics of the gauge fields and their self interactions are described by

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}(W^{i\,\mu\nu} W^i_{\mu\nu} + B^{\mu\nu} B_{\mu\nu} + G^{a\,\mu\nu} G^a_{\mu\nu}) \qquad (1.3)$$

with

$$\begin{aligned}
W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\varepsilon^{ijk} W^j W^k \\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\
G^a_{\mu\nu} &= \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu,
\end{aligned}$$

where ε^{ijk} and f^{abc} are the structure constants[5] of the $SU(2)$ and the $SU(3)$ gauge groups, respectively.
The physical fields of the electroweak gauge bosons A_μ (photon), Z_μ, W_μ^+ and W_μ^- are linear combinations of the W and B fields resulting from the the spontaneous symmetry breaking discussed in the next section with

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2) \qquad (1.4)$$

$$A_\mu = \cos\theta_w B_\mu + \sin\theta_w W_\mu^3 \qquad (1.5)$$

$$Z_\mu = \cos\theta_w W_\mu^3 - \sin\theta_w B_\mu, \qquad (1.6)$$

where θ_w is the Weinberg angle (also called weak mixing angle) defined by $\cos\theta_w = g/\sqrt{g^2 + g'^2}$.

[5] The structure constants $x^{\alpha\beta\gamma}$ are defined by $[m_\alpha, m_\beta] = i x^{\alpha\beta\gamma} m_\gamma$ where m_i are the generating matrices of the symmetry group and $[\dots]$ is the anti-commutator.

1.2.3 The Higgs Mechanism

Due to the required $SU(2)_L$ invariance of the Lagrange density, it cannot contain terms of the form $\overline{\psi}\psi$ [6]. This means that Dirac mass terms of the form $m\overline{\psi}\psi$ are not allowed in the Lagrange density. Therefore, the mass of elementary particles cannot be a generic feature.
To be able to give particles in the SM their observed mass, an additional weak isospin doublet of complex scalar fields ϕ can be introduced [139, 164, 169]. The dynamics of these fields is given by the term in the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = \left|(i\partial_\mu - g\tau^i W_\mu^i - gY'_W B_\mu)\phi\right|^2 + V(|\phi|), \tag{1.7}$$

which also describes the interaction between these scalar fields and the electroweak gauge fields. The idea is that the potential, $V(|\phi|)$, of the scalar field does not have its minimum at $\phi = 0$ [7]. This means that the vacuum state is not symmetric with respect to the potential, leading to a *spontaneous symmetry breaking* of the $SU(2)_L \times U(1)_Y$ invariance to the residual $U(1)_Q$ symmetry of Quantum Electro Dynamics (QED). The resulting non-symmetric ground state of the scalar field doublet can be defined as

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{1.8}$$

where v is the so-called vacuum expectation value. Due to this non-zero ground state, the Lagrange density gets mass terms for the massive gauge bosons without breaking the gauge invariance of the theory. The resulting masses of the Z and W bosons have the relation $m_W = \cos\theta_w m_Z$. In addition this doublet can form with the fermion doublets and singlets mass terms for the fermions

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= U_R^i \overline{\phi} Y_u^{ij} Q_L^j + D_R^i \overline{\phi} Y_d^{ij} Q_L^j + \text{h.c.} \\ &+ E_R^i \overline{\phi^c} Y_l^{ij} L_L^j + \text{h.c.}\end{aligned} \tag{1.9}$$

Here, $Y_{u,d,l}^{ij}$ are the coupling constants, called Yukawa couplings, of the up- and down-type quarks as well as the charged leptons to the Higgs fields ϕ. The fermion masses are therefore proportional to their couplings to this field.
So far, the SM Lagrange density has been degenerated across the three families of quarks and leptons. But the quarks and the charged leptons differ in mass, which means that the Yukawa couplings are not scalars, but rather 3×3-matrices.
In the case of the quarks, the two matrices $Y_{u,d}^{ij}$ require four independent unitary matrices V_{dL}, V_{uL}, V_{uR}, V_{dR} to be diagonalised

$$Y_{u,\text{diag}} = V_{uL} Y_u V_{uR}^\dagger \tag{1.10}$$

$$Y_{d,\text{diag}} = V_{dL} Y_d V_{dR}^\dagger, \tag{1.11}$$

[6] ψ can be split into a left- and right-handed part $\psi = \psi_L + \psi_R$. Further $\overline{\psi}_L \psi_L = \overline{\psi}_R \psi_R = 0$. So

$$\overline{\psi}\psi = \overline{\psi}_R \psi_L + \overline{\psi}_L \psi_R,$$

which makes it obvious that the left-hand side of the equation above is not invariant under a $SU(2)_L$ transformation.

[7] Such a potential needs to have terms of order 4. The potential could also contain terms of $|\phi|$ higher than order 4, but this would get the Higgs theory non-renormalisable [163].

but only three of them can be chosen freely. Therefore, the Yukawa matrices of the up- and down-type quarks are not simultaneously diagonalisable having the consequence that the mass and flavour eigenstates of one type (up or down) of quarks are not identical. By convention, the down-type quarks are chosen and the unitary 3×3 CKM-matrix V_{CKM} [8] [100, 180] is introduced, connecting the mass eigenstates $|d\rangle$, $|s\rangle$, $|b\rangle$ and the flavour eigenstates $|d'\rangle$, $|s'\rangle$, $|b'\rangle$

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}. \tag{1.12}$$

The CKM-matrix given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.13}$$

can be described by three real parameters and a complex phase explaining the origin of the CP *violation* in the electroweak quark sector.
The elements of the CKM matrix also describe the relative transition amplitudes in the weak interaction between different quark flavours. The matrix has almost a diagonal shape, where $|V_{ud}|$, $|V_{cs}|$, $|V_{tb}|$ are of the order one while $|V_{us}|$ and $|V_{cd}|$ are about one, $|V_{ts}|$ and $|V_{cb}|$ two and $|V_{td}|$ and $|V_{ub}|$ three orders of magnitude smaller. (The exact experimental values can be found in Ref. [81].) Hence, transitions across quark families, especially those between the first and the third families are heavily suppressed. This suppression is called *Cabbibo suppression*.

1.3 Flavour Physics

As explained in Sec. 1.1.1, quarks are confined in colour-less hadrons and build composite objects. The difference in the number of quarks and anti-quarks must be a multiple of three due to the structure of the strong interaction. Hadrons made of a quark-anti-quark pair are called *mesons* while states formed by three (anti-)quarks are called *(anti-)baryons*. Other possible states made for example of four or five (anti-)quarks have not yet been observed[9].
The study of hadrons as composite objects made of quarks can lead to a deep understanding of the flavour structure of the SM and can serve as a tool to search indirectly for signs of physics beyond the SM. For example a significant deviation of the measured CKM matrix elements from the predicted unitary behaviour would point to the existence of a flavour structure beyond the SM, *e.g.* a fourth quark family.

1.3.1 b Hadrons

Especially interesting for studying the flavour structure of the SM are hadrons containing heavy quarks, *i.e.* charm and beauty quarks[10], as they offer a broad range of decay channels, which

[8] CKM stands for Cabibbo, Kobayashi and Masukawa. Nicola Cabibbo described the formalism for two quark families while Makoto Kobayashi and Toshihide Masukawa extended it to three families. The latter two have been presented the Nobel Prize in Physics 2008 "for the discovery of the origin of the broken symmetry which predicts the existence of at least three families of quarks in nature".

[9] A recent measurement of the exotic $Z(4430)^-$ state by LHCb suggests the existence of tetra-quark states [32].

[10] Bound states with top quarks have been so far not observed and are also not expected to exist. The reason is that the decay width of the top quark is much larger than the Λ_{QCD} parameter governing the timescale of the

allow to measure precisely the several elements of the CKM matrix as well as to study CP violation and rare decays as potential areas to detect signatures of contributions from physics beyond the SM.
The ground states of hadrons containing a b or $\bar{b}$ quark have relatively long lifetimes of the order of 1 ps due to the heavier mass of the top quark forbidding a decay inside the same quark family and the Cabibbo suppression of decays to the other families. Table 1.4 summarises the properties of the b hadrons relevant for the work described in this thesis.

Table 1.4: Properties of the b hadrons relevant in this thesis. If not otherwise noted the values are taken from Ref. [81].

Hadron	**State**	**Mass m [MeV/c^2]**	**Mean lifetime τ [ps]**
B^+	$\|\bar{b}u\rangle$	(5279.26±0.17)	(1.641±0.008)
B^0	$\|\bar{b}d\rangle$	(5279.58±0.17)	(1.519±0.007)
B_s^0	$\|\bar{b}s\rangle$	(5366.77±0.23)	(1.463±0.032)
B_c^+	$\|\bar{b}c\rangle$	(6276.3 ±1.8) [27]	(0.452±0.033)
Λ_b^0	$\|bud\rangle$	(5619.4 ±0.6)	(1.429±0.024)

1.3.2 Heavy-Quark Effective Field Theory and Operator Product Expansion

Processes in Flavour Physics involving heavy quarks are running due to their heavy masses at an energy scale much larger than the QCD scale $\Lambda_{\mathrm{QCD}} \ll m_{c,b}$ [11]. On the other hand, they run at an energy scale much smaller than the mass of the virtual particles (W and Z bosons) that are involved in the processes.
A full Quantum Field Theoretical calculation of these processes would be the most precise way to describe them, but currently poses a unsurmountable challenge due to the different mass scales that are involved. Effective Field Theories (EFTs) offer a general theoretical framework to reduce the calculation to separate problems for the different scales involved in the process[12].
Heavy-Quark Effective Field Theory (HQEFT) and the Operator Product Expansion (OPE) are based on the premise that heavy flavour processes can be factorised into high energy (short distance) contributions, which can be integrated out and be absorbed by couplings described by Wilson coefficients C_i, and low energy (large distance) contributions from the dynamics of the hadrons, described by operators $\mathcal{O}_i$. The amplitude of a process between an initial state i and final state f can therefore be described by an effective Hamiltonian

$$\mathcal{M}_{i\to f} \quad = \quad \langle f|\mathcal{H}_{\mathrm{eff}}|i\rangle \tag{1.14}$$

given by

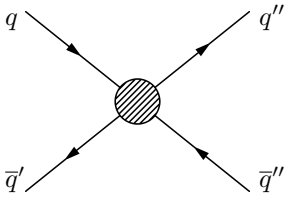
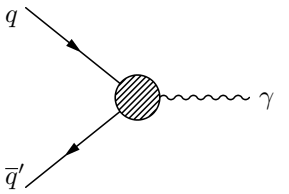
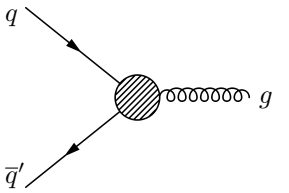
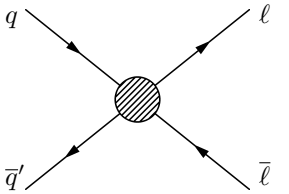
$$\mathcal{H}_{\mathrm{eff}} = \frac{4G_F}{\sqrt{2}} \sum_i \lambda_i^{\mathrm{CKM}} C_i(\mu)\mathcal{O}_i(\mu), \tag{1.15}$$

hadronisation.

[11] For the $\overline{\mathrm{MS}}$ scheme Λ_{QCD} is about 200 MeV [98].

[12] The most prominent is the Fermi Theory of the beta decay, which is an EFT of the electroweak theory, approximating it well for energies far smaller than the mass of the W bosons [145].

Table 1.5: Summary of operators in an Effective Field Theory contributing to a transition of type $q \to q'$.

Operator		**Type**	**Diagram**
$\mathcal{O}_1$	$\propto (\overline{q}'\gamma_\mu T^a P_L q'')(\overline{q}''\gamma^\mu T^a P_L q)$	Tree process	
$\mathcal{O}_2$	$\propto (\overline{q}'\gamma_\mu P_L q'')(\overline{q}''\gamma^\mu P_L q)$		
$\mathcal{O}_3$	$\propto (\overline{q}'\gamma_\mu P_L q)(\overline{q}''\gamma^\mu q'')$	Gluonic Penguin	
$\mathcal{O}_4$	$\propto (\overline{q}'\gamma_\mu T^a P_L q)(\overline{q}''\gamma^\mu T^a q'')$		
$\mathcal{O}_5$	$\propto (\overline{q}'\gamma_{\mu_1}\gamma_{\mu_2}\gamma_{\mu_3} P_L q)(\overline{q}''\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3} q'')$		
$\mathcal{O}_6$	$\propto (\overline{q}'\gamma_{\mu_1}\gamma_{\mu_2}\gamma_{\mu_3} T^a P_L q)(\overline{q}''\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_3} T^a q'')$		
$\mathcal{O}_7$	$\propto e m_q(\overline{q}'\sigma^{\mu\nu} P_R q) F_{\mu\nu}$	Photonic Penguin	
$\mathcal{O}_8$	$\propto g_s m_q(\overline{q}'\sigma^{\mu\nu} T^a P_R q) G^a_{\mu\nu}$	Gluonic Penguin	
$\mathcal{O}_9$ ($\mathcal{O}_V$)	$\propto (\overline{q}'\gamma_\mu P_L q)(\overline{\ell}\gamma^\mu \ell)$	Vector Penguin	
$\mathcal{O}_{10}$ ($\mathcal{O}_A$)	$\propto (\overline{q}'\gamma_\mu P_L q)(\overline{\ell}\gamma^\mu\gamma_5 \ell)$	Axial-vector Penguin	
$\mathcal{O}_S$	$\propto (\overline{q}' P_R q)(\overline{\ell}\ell)$	Scalar Penguin	
$\mathcal{O}_P$	$\propto (\overline{q}' P_R q)(\overline{\ell}\gamma_5 \ell)$	Pseudo-scalar Penguin	

where G_F is the Fermi constant, λ_i^{CKM} describes the flavour structure of the process and μ is the scale separating the short and long distance regimes. The latter is typically chosen as the mass scale of the involved heavy quark.
Table 1.5 lists the different operators that describe the transition between two quark flavours q and q' [79, 117].

Not all operators can contribute to every decay as they differ in the final state. Each operator has a corresponding operator $\mathcal{O}'_i$ with opposite helicity structure, *i.e.* $P_L \leftrightarrow P_R$ where $P_{L,R} = (1 \mp \gamma_5)/2$ are the helicity operators. In the SM, the operators O_{1-10} are allowed and the operators O'_{1-10} are suppressed by a factor $m_{q'}/m_q$ due to the helicity structure of the SM. The operators $O^{(\prime)}_{S,P}$ are forbidden in the SM due to the lack of (pseudo-)scalar couplings in the electroweak

theory[13].
The approach of OPE will be used in Chapter 6 to describe the process of $b \to s\bar{\ell}\ell$, which is relevant for the decay $B^0_{(s)} \to \mu^+\mu^-$.

1.4 b Production at Hadron Colliders

There are three relevant production mechanisms that contribute to the production of $b\bar{b}$ quark pairs at hadron colliders[14]. Figures 1.1 and 1.2 show the Feynman diagrams of the different mechanisms. At second order of the strong interaction coupling strength ($\mathcal{O}(\alpha_s^2)$) there are the processes $q\bar{q} \to b\bar{b}$ and $gg \to b\bar{b}$ summarised as *pair production* [122]. At third order of α_s the next-to-leading order (NLO) corrections to the pair production [201] as well as new contributions of the form $q\bar{q} \to b\bar{b}g$, $gg \to b\bar{b}g$, $gq \to b\bar{b}q$ and $g\bar{q} \to b\bar{b}\bar{q}$ take place. If a heavy sea-quark from one of the beam particles is scattered on a parton from the other beam particle in these $2 \to 3$ processes, the mechanism is called *flavour excitation*[15]. Otherwise, a $g \to b\bar{b}$ branching occurs and the mechanism is classified as *gluon splitting*. At higher order of α_s there are additional diagrams and higher-order corrections contributing to the heavy quark production, but these are typically negligible at the energies reached so far at hadron colliders [205].
The total cross section for the production of $b\bar{b}$ pairs depends on the centre-of-mass energy, $\sqrt{s}$, of the hadron-hadron collision. Figure 1.3 shows different relevant cross sections in pp collisions as a function of $\sqrt{s}$. The $b\bar{b}$ cross section shows approximately a linear dependence on $\sqrt{s}$ in the centre-of-mass energy region of the LHC.
Figure 1.4 illustrates that the relative contribution of the three mechanisms to heavy flavour production also depends on $\sqrt{s}$. As expected, the relative contribution from higher-order processes is increasing with the centre-of-mass energy. For the energies at the LHC the dominating mechanism is flavour excitation.
Table 1.6 shows predicted and measured production cross sections for $b\bar{b}$ production, $\sigma_{b\bar{b}}$, at the LHC-relevant values of $\sqrt{s}$ as well as the fraction of b quarks produced in the acceptance of the LHCb experiment ($2 < \eta < 5$ with $\eta = -\ln[\tan(\theta/2)]$ and where θ is the polar angle with respect to the beam axis.). The calculated values are based on leading order (LO) calculations using Pythia8 [217] with the MSTW08 LO parton distribution function set [195].
At values of $\sqrt{s}$ reached at the LHC, b quarks are produced predominantly with small azimuthal angles θ with respect to the beam axis, as there is typically a significant asymmetry in the proton momentum fractions x_1 and x_2 carried by the two partons taking part in the scatter. Figure 1.5 illustrates this by showing the θ and η dependencies of the b quark production in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. These dependencies have been relevant for the design of the LHCb experiment, which will be discussed in Chapter 3.
Besides the direct production of $b\bar{b}$ quark pairs a negligible amount of b and $\bar{b}$ quarks originates from decays of top quarks, W, Z and Higgs bosons.

[13] An exception is the SM Higgs, but its contribution is negligible due to the small Yukawa couplings.
[14] The qualitative statements in this subsection are also valid for c quarks.
[15] Flavour excitation of quarks other than valence quarks is only possible at $\mathcal{O}(\alpha_s^3)$.

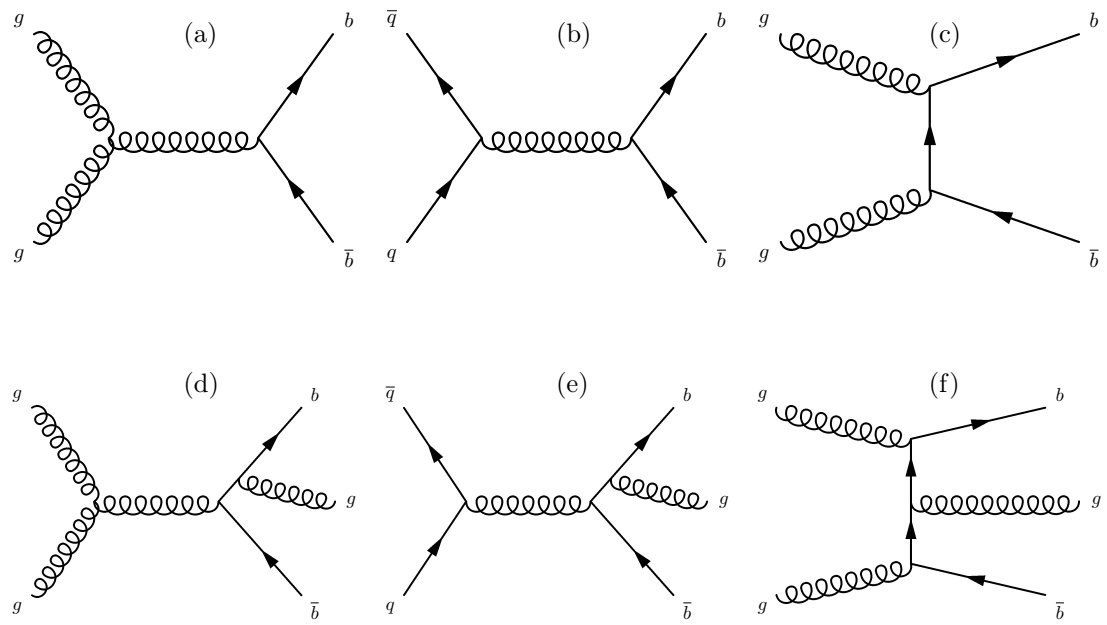


Figure 1.1: Feynman diagrams of $b\bar{b}$ pair production at LO ($\mathcal{O}(\alpha_s^2)$) (a-c) as well as of their NLO corrections ($\mathcal{O}(\alpha_s^3)$) (d-f).

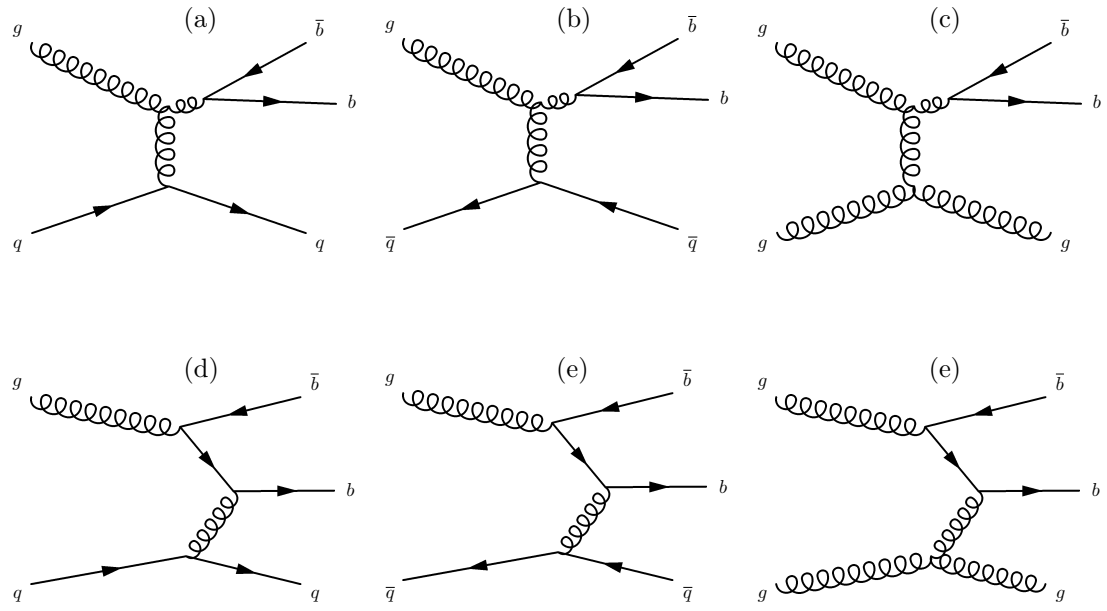


Figure 1.2: Feynman diagrams for $b\bar{b}$ gluon splitting (a-c) and flavour excitation (d-f) at LO ($\mathcal{O}(\alpha_s^3)$).

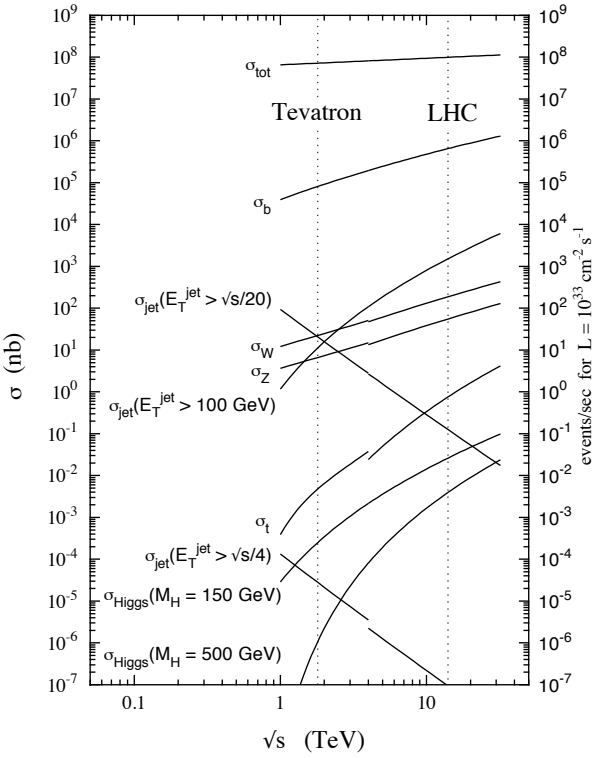


Figure 1.3: Cross sections as a function of the centre-of-mass energy for relevant processes in $p\overline{p}$ colliders (left part: TeVatron) and pp colliders (right part: LHC). The $b\overline{b}$ cross section has approximately a linear dependence on $\sqrt{s}$ in the centre-of-mass energy region of the LHC. (from Ref. [104]).

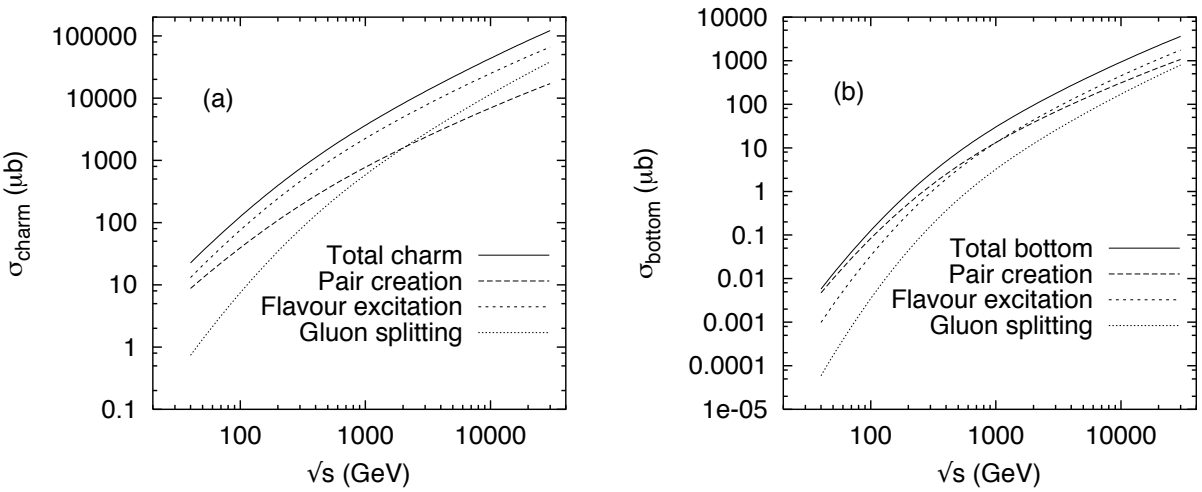


Figure 1.4: Contribution of the three production mechanisms Pair Production, Flavour Excitation and Gluon Splitting as a function of the centre-of-mass energy for (a) $c\overline{c}$ and (b) $b\overline{b}$ production. (from Ref. [205]).

Table 1.6: $b\bar{b}$ cross section in pp collisions, measured by LHCb and predicted by PYTHIA8 in LO calculations. The first uncertainty quoted for the measured values is statistical, the second is systematic.

centre-of-mass energy	**LHCb measurement (LHCb acceptance)**	**LHCb measurement (extrapolated to full phase space)**	**LO calculation (PYTHIA8) (full phase space)**
7 TeV	(75.3±5.4±13.0) µb [18]	(284±20±49) µb [18]	(271.8±0.3) µb
8 TeV	–	–	(319.5±0.3) µb
14 TeV	–	–	(619.0±0.5) µb

centre-of-mass energy	**Fraction of $b\bar{b}$ pairs in the LHCb acceptance**	**Fraction of b and $\bar{b}$ quarks in the LHCb acceptance**
7 TeV	14.81 %	25.59 %
8 TeV	14.80 %	25.59 %
14 TeV	14.58 %	25.33 %

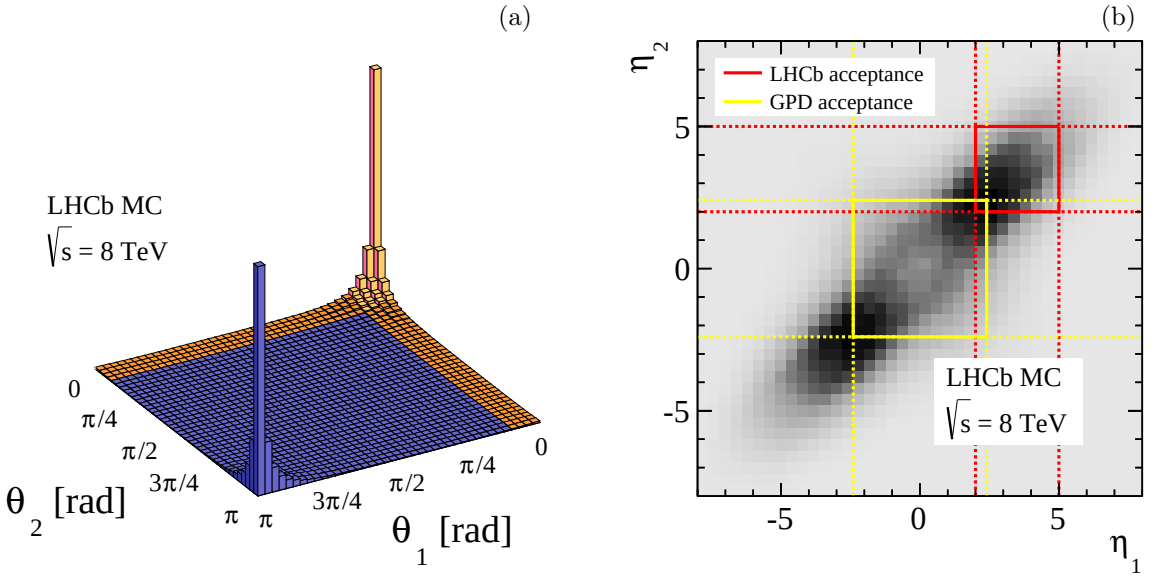


Figure 1.5: Distribution of (a) the azimuthal angles θ and (b) the pseudorapidity η for the b (θ_1, η_1) and the $\bar{b}$ (θ_2, η_2) quarks produced in pp collisions at a centre-of-mass energy $\sqrt{s} = 8$ TeV. The bright shaded region in (a) and the red lines in (b) show the acceptance of LHCb, the yellow lines in (b) that of CMS and ATLAS.

2

CERN and the LHC

This chapter gives a short overview of the history of the Conseil Européen pour la Recherche Nucléaire (CERN) and of the Large Hadron Collider (LHC) with its experiments.

The *Conseil Européen pour la Recherche Nucléaire* (CERN) was established on a provisional basis in 1952 by eleven European countries with the goal of founding a world-class fundamental physics research organisation for Europe. After selecting Geneva as the site for its laboratory in 1952, its now twelve founding member states[1] established the research organisation in 1954[2] [1]. In the same year the construction work for the laboratory began and in 1957 the 600 MeV Synchrocyclotron (SC) started operation as the first accelerator at CERN. In the following years there has been the building and operation of several further accelerator representing milestones for particle physics such as the 28 GeV Proton Synchrotron (PS), the Intersecting Storage Rings (ISR; first proton-proton-collisions), the Super Proton Synchrotron (SPS; used as proton-anti-proton collider to discover the Z and W bosons) and the Large Electron-Positron Collider (LEP). Some of these accelerators are still used and serve as pre-accelerators for the LHC (*cf.* following section) [237].

2.1 The LHC

The *Large Hadron Collider* is a two-ring superconducting hadron accelerator and collider, which has been installed in a pre-existing tunnel used between 1989 and 2000 for the LEP collider. The tunnel has a circumference of 26.7 km and is located in the franco-swiss border region between Lac Léman and the Jura mountains.
The accelerator was designed to operate at a center-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ and each beam consists of a maximum of $n_b = 2808$ bunches, corresponding to a nominal bunch spacing of 25 ns. The maximum design bunch intensity is $N_b = 1.15 \times 10^{11}$ protons per bunch leading

[1] They have been Belgium, Denmark, France, the Federal Republic of Germany, Greece, Italy, the Netherlands, Norway, Sweden, Switzerland, the United Kingdom, and – as the only Eastern European country – Yugoslavia.
[2] Although the provisional CERN has been dissolved, its acronym remained.

to a total beam current of 0.584 A and a stored energy of 362 MJ in the two beams[3] [142].
As the main goal of the LHC is to measure SM processes involving heavy particles as well as processes with small cross sections σ and to reveal physics beyond the SM, a high enough event rate R requires a large instantaneous luminosity of the machine $\mathcal{L}$ as

$$R = \sigma\mathcal{L}. \tag{2.1}$$

Assuming Gaussian beam profiles, $\mathcal{L}$ is given by

$$\mathcal{L} = \frac{N_b^2 f \gamma n_b}{4\pi\varepsilon_n \beta^*} F, \tag{2.2}$$

where f is the revolution frequency of the circulating bunches, γ the Lorentz factor, ε_n the normalised transverse beam emittance, β^* the beta function at the interaction point (IP)[4] and F a reduction factor due to the crossing angle of the beams. F depends besides on the crossing angle θ_c also on the longitudinal (σ_L) and transversal (σ_T) RMS of the particle distribution in the bunches,

$$F = 1 \Big/ \sqrt{1 + \left(\frac{\theta_c \sigma_L}{2\sigma_T}\right)^2}. \tag{2.3}$$

The revolution frequency at the LHC is $f = 11.2$ kHz. Combined with n_b a maximal average crossing rate of bunch pairs at the IPs of 31.5 MHz is obtained.
The peak design instantaneous luminosity at the LHC is 10^{34} s^{-1}cm^{-2}, requiring very low β^* values at the corresponding IPs of 0.3 m for the maximal acceptable transverse beam emittance of $\varepsilon_n = 3.75$ µm.
The maximum beam energy reachable at the LHC is limited by the magnetic field integrated along the beam line as a higher beam energy requires according to the Lorentz force a higher magnetic field to keep the particles on a circle with the same radius. The peak dipole field is 8.33 T. This field is generated by superconducting magnets using Niob-Titanium conductors, which are cooled by superfluid helium to temperatures below 2 K. The limiting factor on the magnetic field strength is actually the heat load inside the magnets. Therefore beam losses during operations must be kept minimal as the lost particles increase the temperature in the magnets.

The LHC is not fully circular, but has eight straight approximately 528 m long sections interleaved with arcs. The straight sections serve as locations for experimental or utility insertions. The numbering scheme of the sections is displayed in Fig. 2.1(b). The purposes of the insertions are:

[3] This requires a high beam intensity, which makes it impossible to use anti-protons for one of the beams.
[4] The beta function at the IP can be interpreted as the distance along the beam between the interaction point and the point where the beam width is twice as large as at the IP. Hence, it is a measure for the beam focussing at the interaction points.

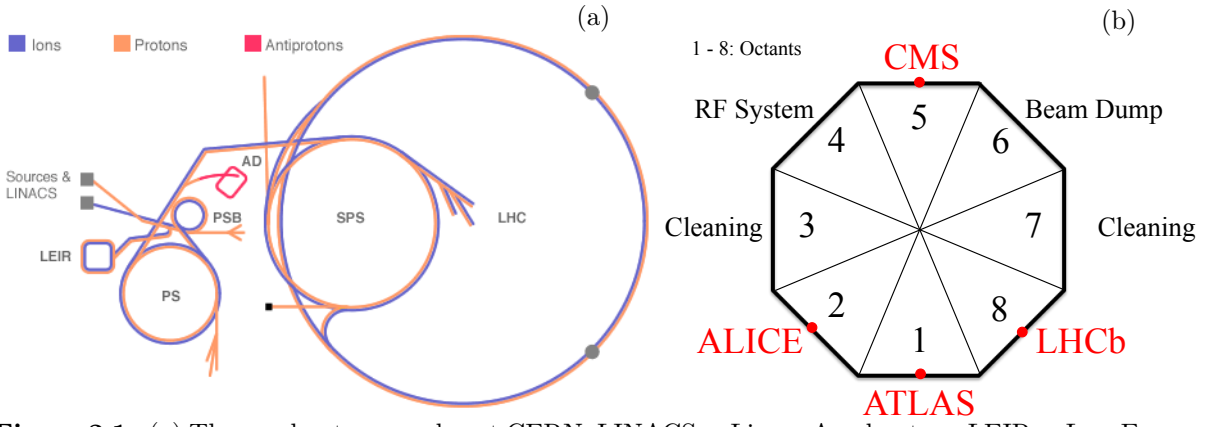


Figure 2.1: (a) The accelerator complex at CERN: LINACS = Linear Accelerators, LEIR = Low Energy Ion Ring, PS = Proton Synchrotron, PSB = Proton Synchrotron Booster, AD = Anti-proton Decelerator, SPS = Super Proton Synchrotron, LHC = Large Hadron Collider (from Ref. [142]). (b) Scheme of the eight sections in the LHC and their purposes.

Section	**Purpose**
1	ATLAS experiment
2	ALICE experiment
3	Collimation System (Momentum)
4	Radio Frequency (RF) systems
5	CMS experiment
6	Beam dump lines
7	Collimation System (Betatron)
8	LHCb experiment

The collimation systems in Sections 3 and 7 have the task to correct the trajectories of particles with high momentum offsets (Section 3) or high betatron amplitudes (Section 7) in the beams by collimating quadrupole magnets.

There are two different Radio Frequency (RF) systems in Section 4 (one per beam) serving as main accelerating system for the beams. They are based on niobium sputtered cavities providing longitudinal oscillating electric fields with an operating frequency of approximately 400 MHz. This frequency needs to be precisely tuned to the revolution frequency of the protons in the LHC.

In Section 6, there are two beam dump lines – one for each beam – that allow a safe extraction of the beams from the LHC. A series of fast kicker and septum magnets remove the particles out of the LHC machine. They are followed by dilution kicker magnets, which sweep the beam in an 'e' shaped form such that the particles do not hit the graphite target at a single point after a flight distance of about 750 m from the septum magnets .

The proton beams are injected into the LHC via two 2.5 km long transfer lines from the SPS at a beam energy of 450 GeV. The source of the protons is a bottle of hydrogen gas. The injection chain (*cf.* Fig. 2.1(a)) that is followed by the protons starts with the Linear Accelerator 2 (Linac2) accelerating the protons to 50 MeV after the hydrogen atoms have been ionised by a duoplasmatron. They are then injected from the Linac2 into the Proton Synchrotron Booster (PSB), where they reach 1.4 GeV before being injected into the Proton Synchrotron (PS). The

PS is also responsible for building the proton bunches with a spacing of 25 ns before they are extracted into the Super Proton Synchrotron (SPS) after reaching 25 GeV. There they reach the final energy of 450 GeV before being injected into the LHC.
Beside protons also lead nuclei ($^{208}_{82}$Pb) can be filled into the LHC. They can reach a maximal beam energy per nucleon, which is by a factor 0.3942 (= 82/208) lower than the one for protons. LHC has not only performed lead-lead collisions, but also lead-proton collisions with the protons travelling both clock- and anti-clockwise in the LHC.
The lead nuclei are generated by evaporating lead atoms from an enriched source in a special oven and afterwards ionised by travelling through a plasma gas. The injection chain followed by the lead nuclei starts with the Linear Accelerator 3 (Linac3) bringing them up to an energy of 4.2 MeV per nuclei and injecting them into the ion accumulator and from there into the PS. Afterwards the injection chain is the same as for protons.

2.2 The LHC Experiments

There are four interaction points (Point 1, 2, 5, and 8) along the LHC where the two beams can collide.
Point 1 and 5 are designed for collisions with highest instantaneous luminosity and used by the ATLAS and CMS experiments, respectively, while at Point 2 and 8 the beam configuration is set for a medium instantaneous luminosity. Point 2 is used by the ALICE and Point 8 by the LHCb experiment. The LHCb experiment will be discussed in the next chapter.

2.2.1 The ATLAS Experiment

The ATLAS (A Toroidal LHC ApparatuS) experiment is one of the two general purpose detectors at the LHC. The detector is equipped with all its sub-detectors in the pseudorapidity range of $|\eta| < 2.5$. All sub-detectors consist of a barrel part covering the central region $|\eta| \lesssim 1.2$ and endcap parts covering the forward regions $1.2 \lesssim |\eta| \lesssim 2.5$.
The Inner Detector consists of tracking detectors (silicon pixel and micro-strip sensors as well as a transition radiation tracker) placed inside a 2 T magnetic field generated by a toroid magnet. The outer part of the detector consists of an electromagnetic (ECAL) and hadronic calorimeter (HCAL) using liquid argon and scintillating tiles. The outermost sub-detectors are Cathode Strip, Drift Tube and Resistive Plate Chambers to measure muons [10].

2.2.2 The CMS Experiment

The second general purpose detector at LHC is called CMS (Compact Muon Solenoid). It covers as ATLAS the phase space region of $|\eta| < 2.5$. Its core is a 4 T superconducting solenoid magnet. The covered pseudorapidity range can be also split into a central barrel region of $|\eta| \lesssim 1.5$ and endcap regions $1.5 \lesssim |\eta| \lesssim 2.5$.
The tracking system is based completely on silicon pixel and micro-strip sensors. The ECAL uses lead tungsten ($PbWO_4$) crystals read out by avalanche photodiodes and vacuum phototriodes while the HCAL uses brass plates as absorbers and plastic scintillating tiles as active material. In the region of highest occupancy, steel is used instead of brass and quartz fibers are used as active material. All these sub-detectors are placed inside the coil of a superconducting 4 T-solenoid

magnet. The part of the detector outside of the solenoid coil is formed by a residual part of the HCAL to detect large showers leaking out of the calorimeter and the muon stations based on Resistive Plate Chambers and Cathode Strip Chambers interleaved with the return yoke of the magnet [108].

2.2.3 The ALICE Experiment

ALICE (A Large Ion Collider Experiment) has the special purpose to measure heavy ion collisions at the LHC. Therefore, the used detector technology has been chosen to cope with the higher occupancy in these collisions happening at smaller rates.
The experiment recycles the 0.5 T solenoid magnet from the LEP experiment L3. The tracking detectors consist of a silicon vertex detector and a Time-Projection Chamber (TPC) as main tracking devices.
ALICE uses arrays of Ring Imaging Cherenkov (RICH), Time-of-Flight (TOF) and Transition Radiation detectors for particle identification. Further there are two different types of electromagnetic calorimeters, one using lead-tungstate crystals as CMS and the other based on stacks of lead as absorbers and polystyrene as active material. To measure heavy-quark resonances such as J/ψ or Υ particles in the forward direction, ALICE is equipped with special tracking stations around a dipole magnet and muon stations. In addition there are further small detectors with dedicated tasks (*e.g.* triggering or detection of cosmic-rays) [39].

2.2.4 Other Experiments

Besides the four large experiments, there are three smaller experiments at the LHC:

- **TOTEM** (Total Elastic and Diffractive Cross Section Measurement) aims to study elastic and diffractive scattering in the very forward direction ($3.1 < |\eta| < 6.5$). It uses gaseous and silicon detectors placed at a distance of 147 m and 220 m on either side of the CMS interaction point [61].
- **LHCf** (Large Hadron Collider forward) is an experiment to collect calibration data for experiments measuring very-high-energy cosmic-rays. There are detectors placed at a distance of 140 m on either side of the ATLAS interaction point. The detectors are position-sensitive calorimeters based on tungsten and plastic scintillator. The position measurements are done by scintillating fibers and silicon strip sensors [45]. The experiment has only taken data during a short phase immediately after the startup of the LHC at low instantaneous luminosity.
- **MoEDAL** (Monopole and Exotics Detector at the LHC) has the goal to search for highly ionising particles (*e.g.* magnetic monopoles). It consists of plastic plates placed around the interaction point of LHCb. Heavy ionising particles produced in the collisions should leave traces in the plastic material, which can be latter detected by an etching procedure after removing the plates from the interaction point [207].

2.3 LHC Operation

First beams were injected and circulated in fall 2008, but soon afterwards an accident damaging several magnets required repair work interventions and a check and consolidation of several

electrical connections in the accelerator. In fall 2009, the LHC was restarted at a centre-of-mass energy $\sqrt{s} = 900$ GeV. In spring 2010 the first data were taken at $\sqrt{s} = 7\,\text{TeV}$. During 2010 the LHC operated with few bunches filled and at a significant lower instantaneous luminosity than the design one while the luminosity was gradually increased in 2011, still at $\sqrt{s} = 7\,\text{TeV}$. The centre-of-mass energy was increased in 2012 to $\sqrt{s} = 8\,\text{TeV}$ and also the instantaneous luminosity was further enhanced.

At the end of the proton-proton run in 2011, LHC also performed lead-lead collisions at a centre-of-mass energy per nucleon-nucleon pair of $\sqrt{s_{NN}} = 2.76\,\text{TeV}$, while after the 2012 proton-proton run proton-lead collisions at a centre-of-mass energy per proton-nucleon pair of $\sqrt{s_{NN}} = 5\,\text{TeV}$ were performed in January and February 2013.

Afterwards, the LHC went into the so-called long-shutdown 1 (LS1) to allow further consolidation work on electrical connections and improvement on some parts of the experiments. The restart after LS1 will take place at the beginning of 2015 and it is planned that the experiments will take first data at a centre-of-mass energy between 13 and 14 TeV in the first half of 2015.

3

The LHCb Experiment

The LHCb experiment [55, 128] is a high-precision experiment at the LHC, designed to study the physics of b and c quarks and their bound states. Its location is at Point 8 of the LHC.
As b and c quarks are at the LHC enhancedly produced under small polar angles with respect to the beam axis (*cf.* Sec. 1.4) and to achieve a very high precision on the momentum measurement by a long flight distance between the first and last tracking stations, the LHCb detector is set up as a single-arm forward spectrometer using a dipole magnet. A schematic view of the detector is shown in Fig. 3.1.
Its coverage in polar angles ranges from 10 mrad to about 300 mrad in the horizontal bending plane of the dipole magnet and to about 250 mrad in the vertical plane. This corresponds to an approximate coverage in pseudorapidity $\eta = -\ln[\tan(\theta/2)]$ from 2 to 5.
The detector has a length of about 20 m and a height and width of up to 10 m. The coordinate system used in LHCb is a cartesian one, with the z-direction pointing along the LHC beam axis from the interaction point toward the Muon System. The y-direction points upward while the x-direction is oriented such that the three axes form a right-handed coordinate system.
Due to a modification of the LHC beam optics the interaction point lies 11.25 m away from the centre of the cavern. This allows to make a maximum usage of the existing cavern for the LHCb detector.
The average LHCb design instantaneous luminosity at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ was $2 \times 10^{32}\ \text{cm}^{-2}\ \text{s}^{-1}$, which is almost a factor 100 lower than the one for ATLAS and CMS[1]. The reason is that the forward region is dominated by a high particle multiplicity creating high occupancies in the detector elements. A higher instantaneous luminosity would therefore severely constrain the choice of the detector design and the used materials. In addition most of the analyses in LHCb require a very good separation of primary and secondary vertices as a displaced decay vertex is a distinctive feature of b and c hadrons. This separation would be reduced by the presence of as many pp interactions in an event as in ATLAS or CMS.
With an expected $b\bar{b}$ cross section at $\sqrt{s} = 14\,\text{TeV}$ of $\sigma_{b\bar{b}} \approx 500\ \mu\text{b}$ there would be a trillion $b\bar{b}$ pair

[1] The lower luminosity is achieved by a transverse separation of the two beams reducing their overlap. As the beam intensity is decreasing during data taking, the separation is reduced resulting in a stable instantaneous luminosity over the fill. This procedure is called *luminosity leveling* [174].

produced in a canonical year corresponding to 10^7 s of data taking. The average design number of visible interactions[2] per bunch crossing, called μ, would have been 0.4. This configuration would have led to an integrated luminosity per year of 2 fb^{-1}.
During the first year of data taking, LHCb collected a data sample corresponding to 38 pb^{-1} at $\sqrt{s} = 7\,\text{TeV}$ while running with an average μ set to 2.7 due to the much smaller number of bunches in the LHC beams. The number of bunches per beam has been increased during the years 2011 and 2012 to a maximum value of 1380. In this period LHCb ran at an average μ of 1.4 and collected 1 fb^{-1} at $\sqrt{s} = 7\,\text{TeV}$ in 2011 and about 2 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$ in 2012. The integrated luminosity as a function of time is shown in Fig. 3.2(a), and Fig. 3.2(b) shows the peak luminosity per fill as a function of time. Table 3.1 summarises the main running parameters for LHCb. The average data taking efficiency for 2010 to 2012 has been over 90 % (*cf.* Fig 3.3 for details).
The sub-detectors in LHCb can be grouped into four sections: a tracking system using a dipole magnet, two Ring Imaging Cherenkov (RICH) detectors, a calorimeter system including an electromagnetic as well as a hadronic calorimeter, and a Muon System.

Table 3.1: Summary of the running conditions in 2010-2012. [239]

Quantity	**2010**	**2011**	**2012**
Centre-of-mass energy	7 TeV	7 TeV	8 TeV
Nominal number of bunches per beam	368	1381	1374
Nominal number of colliding bunch pairs in LHCb	344	1292	1262
Average instantaneous luminosity in LHCb [cm^{-2} s^{-1}]	1.2×10^{30}	3.5×10^{30}	4.0×10^{30}
Nominal average μ	2.7	1.4	1.7

3.1 The Tracking System

The tracking system in LHCb consists of a Vertex Locator (VELO) and a tracking station upstream of a dipole magnet, called Tracker Turicensis (TT). Downstream of the dipole magnet there are three tracking stations (T stations), called T1-3 and consisting of an inner part close to the beam pipe made of silicon micro-strip sensors (Inner Tracker; IT) and an outer part made of straw tube sensors (Outer Tracker; OT). The TT and the IT together are also referred to as Silicon Tracker (ST).
While the VELO has the goal to measure the primary and secondary vertices as precisely as possible, the aim of the tracking stations is to detect tracks of charged particles and determine their momenta by measuring the bending of their trajectories in the field of the dipole magnet.

3.1.1 The Vertex Locator

The VELO [55, 76] uses silicon micro-strip modules providing measurements of the r- and ϕ-coordinates (radial distance from the beam axis and azimuthal angle in the x-y plane) of track

[2] A visible interaction is defined as at least two tracks reconstructed by the Vertex Locator and the tracking stations downstream of the dipole magnet.

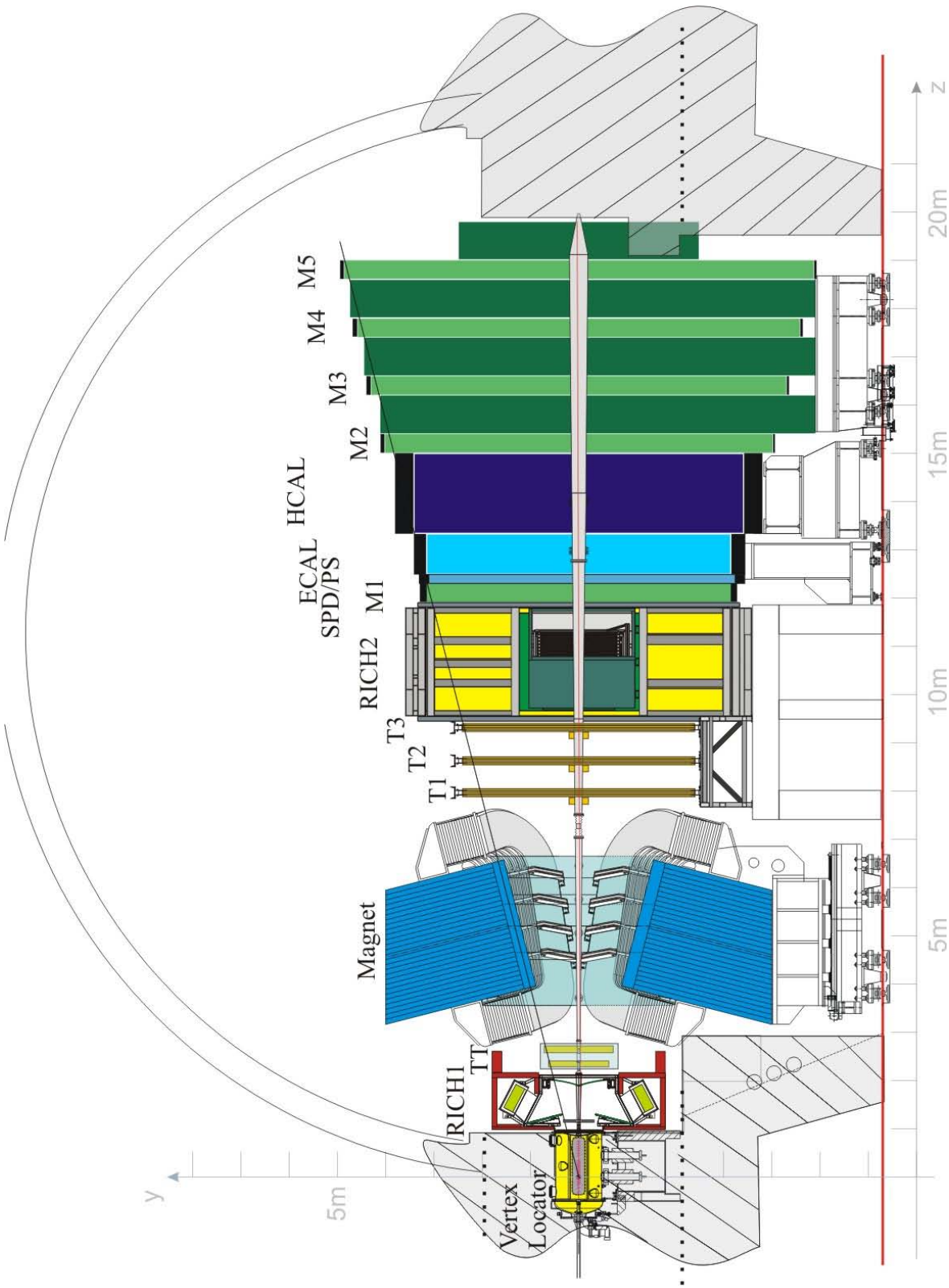


Figure 3.1: Side view of the LHCb detector: RICH1/2: Ring Imaging Cherenkov detectors; TT: Tracker Turicensis; T1-3: Tracking stations; SPD/PS: Scintillating pad detectors/Pre-shower; ECAL: Electromagnetic calorimeter; HCAL: Hadronic calorimeter; M1-5: Stations of the Muon System. (from Ref. [238])

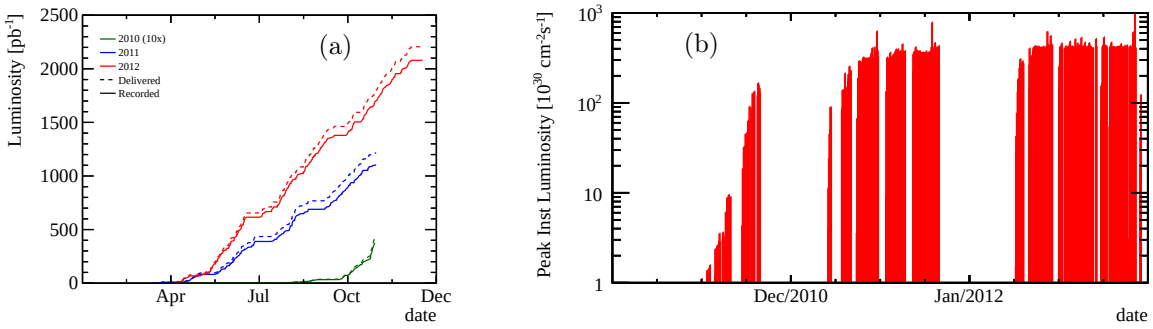


Figure 3.2: (a) Integrated Luminosity in LHCb as a function of time for the years 2010 (green), 2011 (blue) and 2012 (red). The delivered luminosity is displayed by the dashed line, the recorded luminosity by the solid one. (b) Peak instantaneous luminosity per fill as a function of time in the years 2010 to 2012.

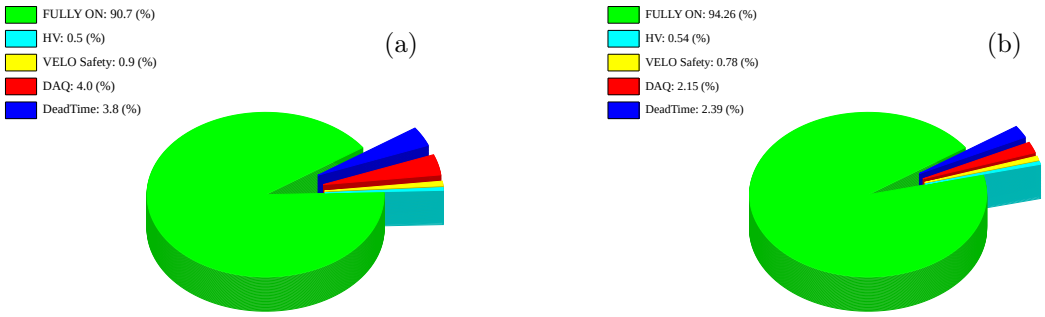


Figure 3.3: The LHCb efficiency breakdown for pp collisions in (a) 2011 and (b) 2012. HV: high voltage settings not OK; VELO Safety: closing of the Vertex Locator at the beginning of the run; DAQ: data acquisition not ready; DeadTime: detector and data acquisition dead time.

hits close to the interaction region. The reconstructed track segments are then used to identify primary and secondary vertices. During the 2012 data taking the impact parameter resolution was about 25 µm for tracks with transverse momenta $p_{\mathrm{T}} = 2\,\mathrm{GeV}/c$ [223].
There are 21 detector modules along the interaction region (*cf.* Fig. 3.4(a)) to measure the track segments in a large pseudorapidity range ($1.6 < \eta < 4.9$) and to have at least three hits in the VELO stations for tracks emerging at $|z| < 11$ cm. Each module consists of two half-disks placed on each side of the beam axis. The modules are designed as single-sided micro-strip silicon sensors paired back-to-back. The strips are on one side arranged such that the r-coordinate is measured while the strips on the other side measure de facto the ϕ-coordinate.
The minimal distance of the modules from the beam axis during data taking is smaller than the aperture required during the injection of the LHC beams. Therefore the VELO can retract the modules during the injection (VELO open) and move them back together during data taking (VELO closed) (*cf.* Fig. 3.4(a)). During data taking the distance of the innermost sensor strips from the beam axis is 8 mm to achieve the best possible impact resolution and the two half-disks

of a module are slightly overlapping. Such a design also requires that the sensors are kept in a secondary vacuum inside the LHC beam vessel, which is only separated from the primary beam vacuum by a 300 µm thin undulated aluminum foil, called RF-foil. This setup ensures a minimal material budget – the maximal distance of a track passing through the VELO is $0.6X_0$ in terms of radiation length X_0 – leading to the above mentioned very good impact parameter resolution.

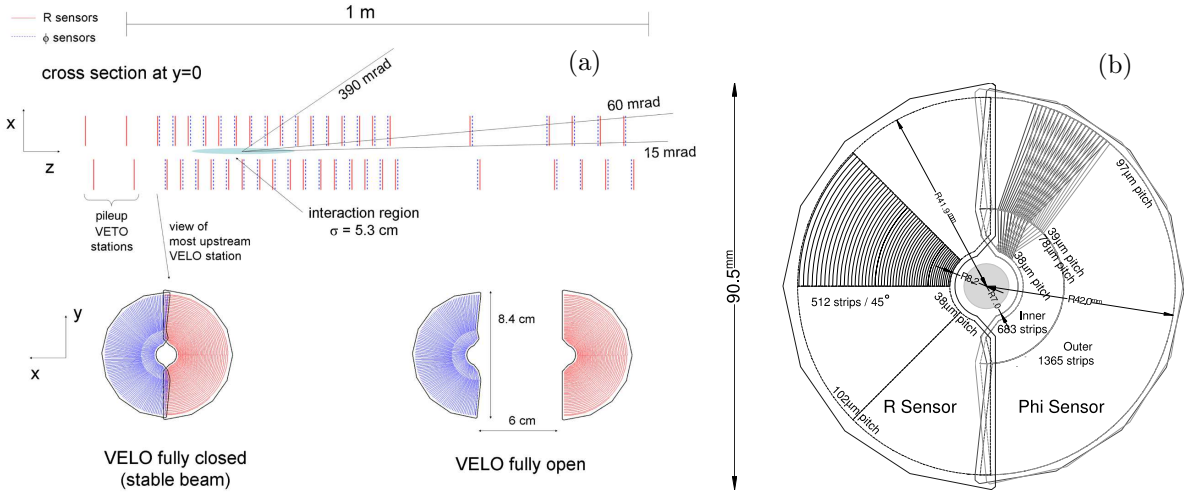


Figure 3.4: (a) Schematic view of the VELO: (top) the layout of the modules along the beam axis and (bottom) the setup of a closed and an open VELO module; (b) Schematic view of a VELO module. The two-half disks are segmented into four (r-sensors) and two (ϕ-sensors) regions, respectively. The numbers indicate the numbers of strips and the strip pitch in different regions of the sensors. (from Ref. [55])

The sensors consist of an n-doped bulk with n^+-implants where the read-out strips are isolated by a p-spray applied in between the strips (there is one module of pure n^+-on-n type placed in one of the most upstream slots). The sensors are 300 µm thick and have a diameter of about 90 mm. The strips of the r- and of the ϕ-sensors are segmented into four and two regions, respectively (*cf.* Fig. 3.4(b)). This segmentation helps to cope with the occupancy in the sensors.
Across the sensors the strip pitch varies due to the circular design of the sensors, but also to ensure that the contributions of the measurements by the different VELO stations along the track to the impact parameter precision have roughly equal weight.
Further the strips in the ϕ-sensors are slightly skewed (in the inner region by 20°, in the outer by 10°). To distinguish fake hits from true hits, adjacent ϕ-sensors have the opposite skew to each other.
The VELO sensors are cooled with bi-phase CO_2 and kept at a temperature of −7°C during operations [46].

3.1.2 The Magnet

The magnet used in the LHCb detector is a warm dipole magnet with an integrated magnetic field of $\int \mathrm{d}lB = 4$ Tm [55, 57]. Its field lines are oriented vertically. The polarity of the field is periodically flipped during data taking to control possible systematic effects on the measurements induced by the magnet polarity.
As the magnetic field is perpendicular to the beam axis, there are corrections in the beam optics

required to take into account the bending of the proton beams by the LHCb magnet.
The yoke of the magnet consists of carbon steel. The two coils are saddle-shaped and are placed mirror-symmetrically above and below the beam pipe in the yoke (*cf.* Fig. 3.5(a) for details). They are made of fifteen layers of highly pure hollow aluminum conductors. The layers can slide freely upon their support to cope with thermal stress.
Figure 3.5(b) shows the measured magnetic field along the z-axis. The field map has been generated based on measurements performed during the construction of the experiment with an array of Hall probes. The achieved relative precision on the magnetic field integral is a few times 10^{-4}. The figure displays also that there are fringe fields upstream and downstream of the magnet. On the one hand the presence of fringe fields between the VELO and TT is beneficial for the tracking. On the other hand the magnetic field must be smaller than 2 mT in the Ring Imaging Cherenkov (RICH) detectors for their photon detectors to function.

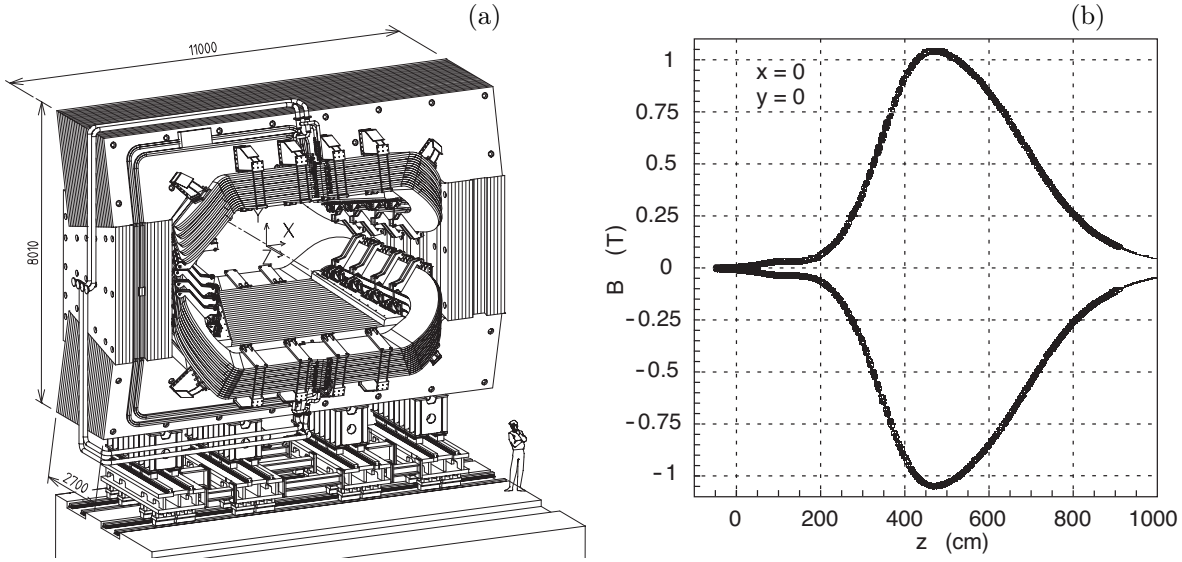


Figure 3.5: (a) Schematic view of the LHCb dipole magnet seen from the downstream region. The dimensions are indicated in millimeters. The frame is the magnet yoke while the saddle shaped devices are the two coils; (b) The main component of the magnetic field B_y as a function of the z-position measured along the beam axis, shown for both polarities. (from Ref. [55])

3.1.3 The Tracker Turicensis

The Tracker Turicensis (TT) [55, 73] forms the tracking station upstream of the dipole magnet. It consists of four layers of silicon micro-strip sensors. The strips in the the first (TTaX) and last layer (TTbX) are arranged vertically while those in the second (TTaU) and third layer (TTbV) are rotated by a stereo angle with respect to the y-axis of $-5°$ and $+5°$, respectively (*cf.* Fig. 3.6). The layers are housed in a large thermally and electrically insulating as well as light tight volume, in which the temperature is kept at about 8°C. The low temperature prevents fast ageing of the sensors due to the irradiation (*cf.* Chapters 4 and 5).
The layers are grouped into two pairs, which have a separation along the beam axis of 27 cm to aid the track reconstruction algorithms. Their basic building blocks are half-modules aligned vertically and covering half the height of the LHCb acceptance. Adjacent modules within a layer

are staggered by about 1 cm in the z-direction and have an overlap of a few millimeters in the x-direction to avoid insensitive areas in the acceptance.
A half-module (*cf.* Fig. 3.7) consists of a row of seven silicon sensors grouped into either two or three read-out sectors and two or three read-out hybrids at the outer end. The half-modules closest to the beam pipe have three sectors consisting of one, two and four sensors (from $y = 0$ to the outer region). The other modules are organised into read-out sectors of three (close to $y = 0$) and four sensors (in the outer region). The read-out strips of the outermost read-out sector are directly bonded to the read-out electronics while those of the inner sectors are connected via Kapton cables.
The silicon sensors are single-sided p^+-on-n type micro-strip sensors made of a negatively (n) doped bulk with an n^+-doped region close to the back plane. On the front side of the bulk there are read-out strips made of aluminum, below which there are positively doped inlets in the bulk. The read-out strips on top of the inlets are decoupled from the silicon bulk by a thin silicon oxide (SiO_2) layer to prevent leakage currents affecting the read-out [240].
The sensors have a thickness of 500 µm and a strip pitch of 183 µm. They are 9.44 cm long and 9.46 cm wide and carry 512 read-out strips. They are operated at a bias voltage of 300 V as they have depletion voltage values between 140 V and 260 V (*cf.* Chapters 4 and 5). In total there are 8.4 m^2 of active area in TT with a total of about 144'000 read-out channels.
The electronic read-out chain starts with the front-end Beetle chip [47] mounted on the read-out hybrid at the outer end of the module. This ASIC chip amplifies, shapes and samples the analogue output signal from the strips. Analogue output signals are transmitted via copper cables with a length up to 8 m to so-called service boxes kept outside of the acceptance, but close to the detector. There the signal is digitalised and subsequently sent via optical fibres using VCSEL[3] diodes to the counting house where the TELL1 [166] boards using FPGAs[4] perform pedestal and common mode noise subtraction as well as zero-suppression on the signal.

3.1.4 The Inner Tracker

The Inner Tracker (IT) [55, 73] forms the inner part of the three tracking stations downstream from the dipole magnet. Each of the three IT stations has four individual detector boxes, one on each side of the beam pipe (A- and C-side boxes), one above (top box) and one below (bottom box) it (*cf.* Fig. 3.8(a)). Each detector box houses four detection layers with seven detector modules each.
The modules consist of two (A- and C-side boxes) or one (top and bottom boxes) silicon micro-strip sensor (*cf.* Fig. 3.8(b)). As in the TT, the modules of the inner two layers in each box are tilted by a stereo angle of $\pm 5^\circ$ while in the outer layers the strips are aligned parallel to the y-axis.
The used sensors are 7.6 cm wide and 11 cm long and carry 384 read-out strips with a strip pitch of 198 µm. They are also designed as single-sided p^+-on-n sensors. The thickness is 320 µm for the sensors used in one-sensor modules and 410 µm for the sensors in two-sensor modules. This ensures a sufficiently high signal-to-noise ratio while keeping the material budget minimal.
In the case of the IT a read-out sector always corresponds to a module and the read-out strips are directly connected to the read-out hybrids placed at one end of the module. The read-out electronics for the IT are identical to the TT with the difference that the front-end hybrids are

[3] Vertical-Cavity Surface-Emitting Laser
[4] Field Programmable Gate Array

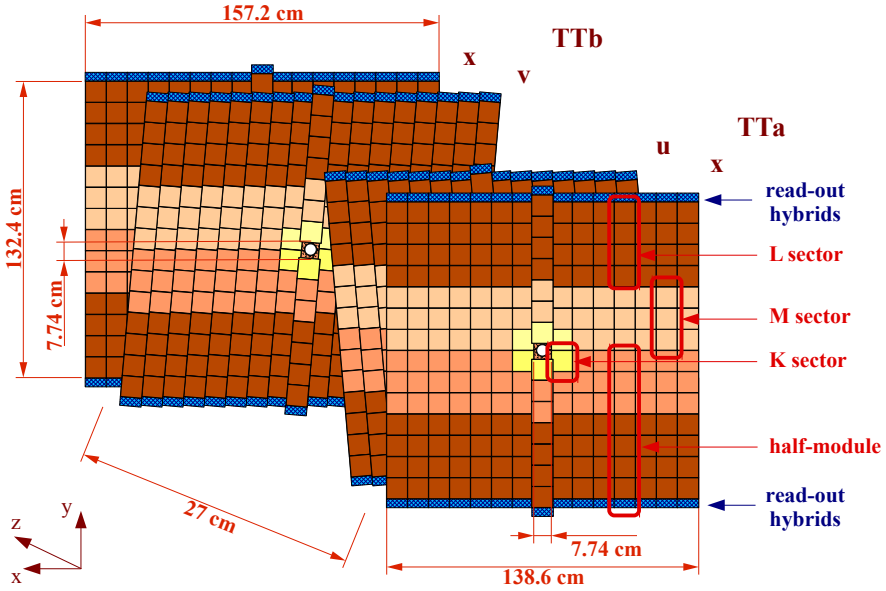


Figure 3.6: The layout of the TT layers seen from the interaction point. (from Ref. [241]). From front to back: TT layer TTaX, TTaU, TTbV, TTbX

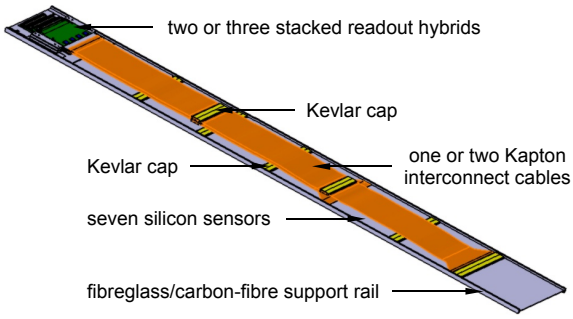


Figure 3.7: The layout of a TT half-module with three read-out sectors. (from Ref. [241])

located inside the detector acceptance.
The IT has an approximate active area of 4 m^2 and in total about 130'000 read-out channels. The IT silicon sensors are also kept at an average temperature of about 8°C.

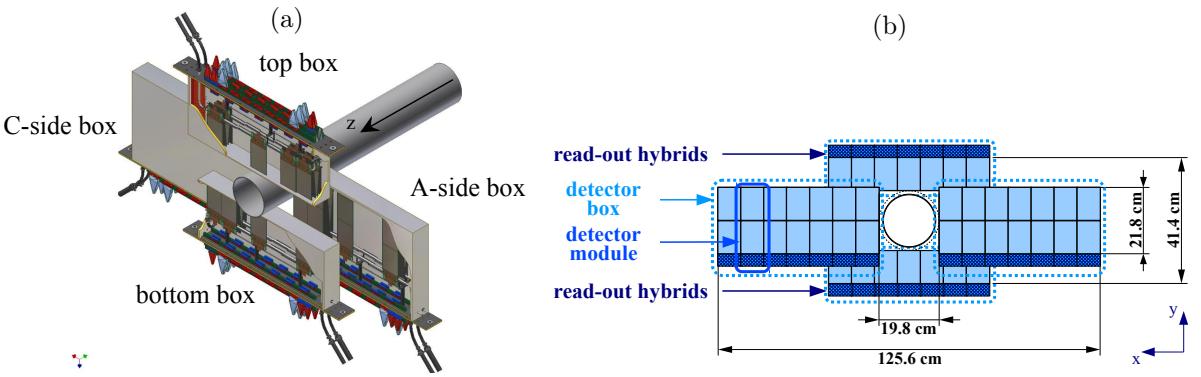


Figure 3.8: (a) The layout of the four detector boxes forming the IT part of one T station around the beam pipe; (b) The layout of one detection layer in the IT with a stereo angle of 0°. (from Ref. [241])

3.1.5 The Outer Tracker

The Outer Tracker (OT) [55,75] forms the outer part of the tracking stations T1-3. It is designed as a drift-time detector made of an array of individual gas-tight straw-tube modules (*cf.* Fig. 3.9). In total there are twelve detector layers (four per station) like in the IT and the stereo angle setup is also the same as in the IT.
There are two different types of modules: long (type F) and short (type S) modules. The first type is used in the region left and right of the beam pipe. The straw tube layers are split at $y = 0$ into two sections, which are read out from each outer end. These modules have an active length of 4850 mm and the two staggered layers in each module consist in total of 256 tubes. The second type of modules consists of 128 straw tubes in two layers and is used above and below the IT. Their length is half that of the F-type modules and they are only read out from the outer end. Each layer is built from eight S-type and fourteen F-type modules leading to a total of about 55'000 single straw-tube channels in OT covering an active area of approximately 29 m^2.
Each module consists of two staggered layers of drift-tubes. Their inner diameter is 4.9 mm and they are filled with a 70/28.5/1.5 %-mixture of Argon, CO_2 and O_2 as counting gas[5] [225]. This gas mixture guarantees a slow enough drift time for a drift-coordinate resolution of 200 µm. The tube is made of an inner (cathode), 40 µm thick foil of electrically conductive Kapton-XC and an outer (laminate) Kapton-aluminum foil forming the actual tube. The anode is a gold-plated tungsten wire with a diameter of 25.4 µm in the center of the tube. It is kept during data taking at a voltage of 1550 V.
The read-out is based on front-end electronic boxes. Each box processes the signal of 128 straw-tubes (*i.e.* one side of an F-type or a full S-type module) whose anode wires are soldered to pads

[5] Originally the mixture was 70/30 % Argon/CO_2, but to prevent ageing effects a small portion of oxygen has been added.

on a circuit board. The front-end electronic measures the drift time of the ionisation clusters. In a first step the signal is decoupled from the high voltage. Afterwards the signal is amplified and shaped before a time-to-digital converter digitises the signal arrival time with respect to the bunch crossing time. The digitised signal is then sent via optical fibres to the TELL1 boards for further processing [82].

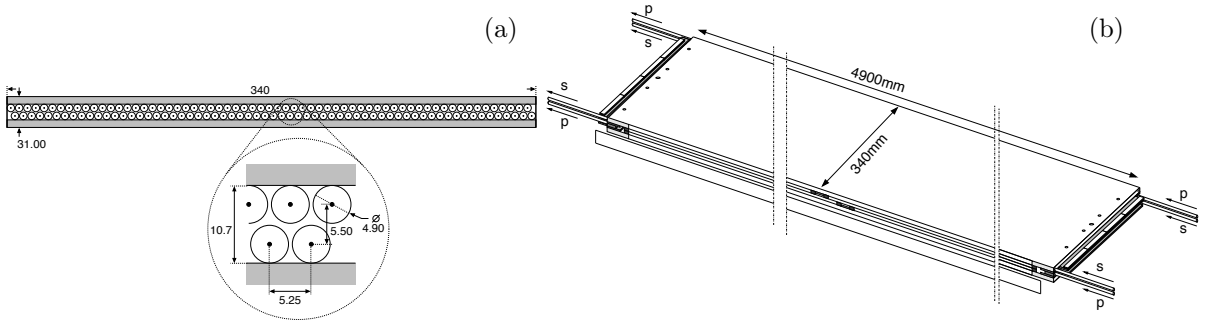


Figure 3.9: The layout of an Outer Tracker module showing (a) the arrangement of the two staggered straw tube rows and (b) the full module with the support structure. (from Ref. [55])

3.1.6 Performance of the Tracking System

The efficiency to find a signal left by a track in a tracking layer is larger than 99 %. For VELO the hit finding efficiency is 99.5 % for a search window of 1 mm around the track [223] while for the ST the efficiency is 99.3 % for a search window of 1 mm [138]. In the case of the OT, the efficiency for detecting a hit of a track crossing the inner half of the radius of a straw tube is 99.2 % [224].

The hit resolution in the VELO depends on the strip pitch, but goes down to 4 µm for a strip pitch of 40 µm [86, 223]. In the case of the TT and the IT, the hit resolution including residual alignment imperfection is 61 µm and 54 µm, respectively. The alignment of the detectors is done with a global χ^2-minimisation of track residuals coming from a Kalman filter fit [60, 172, 173] while in the VELO a non-iterative method with a matrix inversion based on the Millepede procedure is used as a second approach [228].
The relative momentum resolution of the tracking system is $\sigma_p/p = 0.4\,\%$ for Long tracks (*cf.* Chapter 8) with momenta of $p = 5\,\text{GeV}/c$ and increases to 0.6 % for track momenta of $100\,\text{GeV}/c$. The consequent mass resolution is for two-body decays of D^0 mesons from b hadrons $7.2\,\text{MeV}/c^2$ and for the decay $J/\psi \to \mu^+\mu^-$ of J/ψ from b hadrons $12.8\,\text{MeV}/c^2$ (*cf.* Fig. 3.10). The difference in the mass resolution of J/ψ to the numbers shown in Chapter 8 is due to the application of the momentum scale correction and tighter upper cuts on the muon momenta in the sample considered here.

The tracking in LHCb distinguishes between three main track types depending on the subdetectors involved in the track reconstruction:

- **Long tracks:** tracks traversing the LHCb detector from the VELO up to the T stations

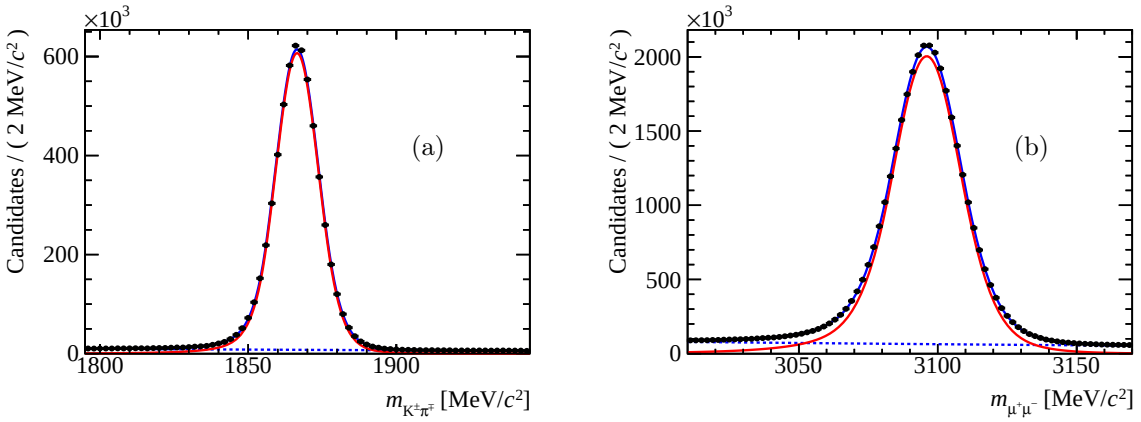


Figure 3.10: Invariant mass distribution of (a) $D^0 \to K^+\pi^-$ and (b) $J/\psi \to \mu^+\mu^-$ candidates in 2012 data. Each of the two distributions is fit by an exponential distribution (blue dotted line) describing combinatorial background and a double Crystal Ball function (*cf.* Eq. (7.1); red solid line) to describe the signal. The total fit model is shown by the blue solid line.

are called Long tracks. They are reconstructed from hits in the VELO and the T stations, but do not necessarily contain hits from the TT. They are the most often used track type in LHCb analyses and are reconstructed by the *Forward Tracking* [103] and the *Track Matching* [202] methods[6].

- **Downstream tracks:** tracks reconstructed only from hits in the TT and the T stations are called Downstream tracks and are typical for charged particles coming from the decay of long living particles (*e.g.* K^0_S, Λ^0). Their reconstruction is done in a similar manner as for the Forward Tracking, but using the TT and T station track segments [102].

- **Upstream tracks:** tracks reconstructed only from hits in the VELO and the TT are named Upstream tracks. They usually come from charged particles with low momenta, which are bent out of the acceptance in the field of the dipole magnet. They are reconstructed by a similar procedure as the Forward Tracking [236].

Further there are special track types for dedicated purposes:

- **VELO tracks:** track segments in the VELO are called VELO tracks. Due to the absence of magnetic field in the VELO these tracks are straight lines with no momentum estimates.

- **T tracks:** track segments in the T stations are called T tracks. The fringe field of the magnet in the region of the T stations leads to a small bending of the track – and therefore a rough momentum estimate.

[6] The Forward Tracking is based on track segments reconstructed in the VELO. They are then matched to hits in the T station by a Hough transform and the selected hits are afterwards fit to a track.
The Track Matching takes track segments reconstructed in the VELO and the T station [148] and matches the segments extrapolated to the bending plane of the dipole magnet.
In both cases hits in the TT are – if possible – added to the track in a second step and a total χ^2-fit is performed. The fit quality serves as a criterion if a reconstructed track candidates is accepted as a track or not.

- **TTMuon tracks:** for the estimation of the reconstruction efficiency for Long tracks, tracks based on hits in the TT station and in the Muon System are reconstructed. These tracks are then matched to Long tracks to estimate the tracking efficiency [31, 126].

The tracking efficiency for Long tracks measured in LHCb from data is for track momenta below 10 GeV/c on average 93 % while it is 97 % on average for track momenta larger than 10 GeV/c [31, 126].

3.2 The Ring Imaging Cherenkov Detectors

There are two Ring Imaging Cherenkov (RICH) detectors in LHCb [55, 58]: a first one, RICH1, upstream of the dipole magnet, between the VELO and the TT, and a second one, RICH2, downstream of the dipole magnet, between the T stations and the calorimeters.
The main purpose of these detectors is the particle identification (PID) of charged hadronic particles (*i.e.* pions (π), kaons (K) and protons (p)), but they also have some PID sensitivity for muons and electrons.
The goal of the RICH detectors is to measure the velocity $\beta = v/c$ of the particles, which can be combined with the momentum measured by the tracking system and transformed into a mass hypothesis for the particle. A charged particle traversing a material faster than the speed of light in it generates Cherenkov radiation. The emission angle with respect to the particle trajectory θ_c is given by

$$\cos\theta_c = \frac{1}{n\beta}, \tag{3.1}$$

where n is the refraction index of the traversed material. So the velocity of the particle can be determined by measuring this emission angle .

3.2.1 RICH1

The RICH1 (*cf.* Fig. 3.11(a)) uses two radiator materials, silica aerogel and fluorobutane gas (C_4F_{10}), and has a π/K separation power for tracks with momenta between 1 and 60 GeV/c. Silica aerogel is a solid form of quartz with a very low density and has a refraction index of 1.03 at a wavelength of $\lambda = 400$ nm. It can provide PID information for low momentum tracks, while the C_4F_{10} has a refraction index of 1.0014 at $\lambda = 400$ nm and gives PID in the range of 10 to 60 GeV/c.
The emitted Cherenkov light is focussed and deflected out of the LHCb acceptance by a combination of flat and spherical mirrors. The photons are detected by position-sensitive Pixel Hybrid Photon Detectors (HPDs) with a wavelength sensitivity between 200 and 600 nm. The incoming photon produces in the HPD a photo-electron in a photocathode, which is then accelerated by a high voltage of the order of 10 kV onto a reverse-biased pixel silicon sensor.
The signal in the silicon sensor is amplified, shaped and compared to a global threshold in a pixel read-out chip. Afterwards the digitised signal is transmitted via optical fibres to electronic boards kept off the detector where the signals are further processed using FPGAs [235].
In RICH1, but also in RICH2, the magnetic shielding to ensure little distortion of the electrons in the HPDs by the fringe field from the dipole magnet is based on 50 mm and 100 mm thick

highly pure ARMCO iron plates. Further the residual distortion is corrected by comparing the measured Cherenkov light spots with the magnet turned on and off.

3.2.2 RICH2

The RICH2 (*cf.* Fig. 3.11(b)) has the same photon detectors and read-out electronics as RICH1. It uses CF_4 gas as radiator material. This radiator has a refractive index of 1.0005 at a wavelength of $\lambda = 400$ nm and is therefore used to perform PID for high momentum particles (*i.e.* track momenta between 16 and 100 GeV/c).
The acceptance of RICH2 extends only out to $\pm$120 (100) mrad in the horizontal (vertical) plane, but covers the region that is passed by the high momentum particles (*cf.* Fig. 3.12).
A design difference to the RICH1 is that the optical system to deflect the Cherenkov radiation is arranged horizontally in RICH2, while in RICH1 the photons are deflected to the top or the bottom of the detector [44].

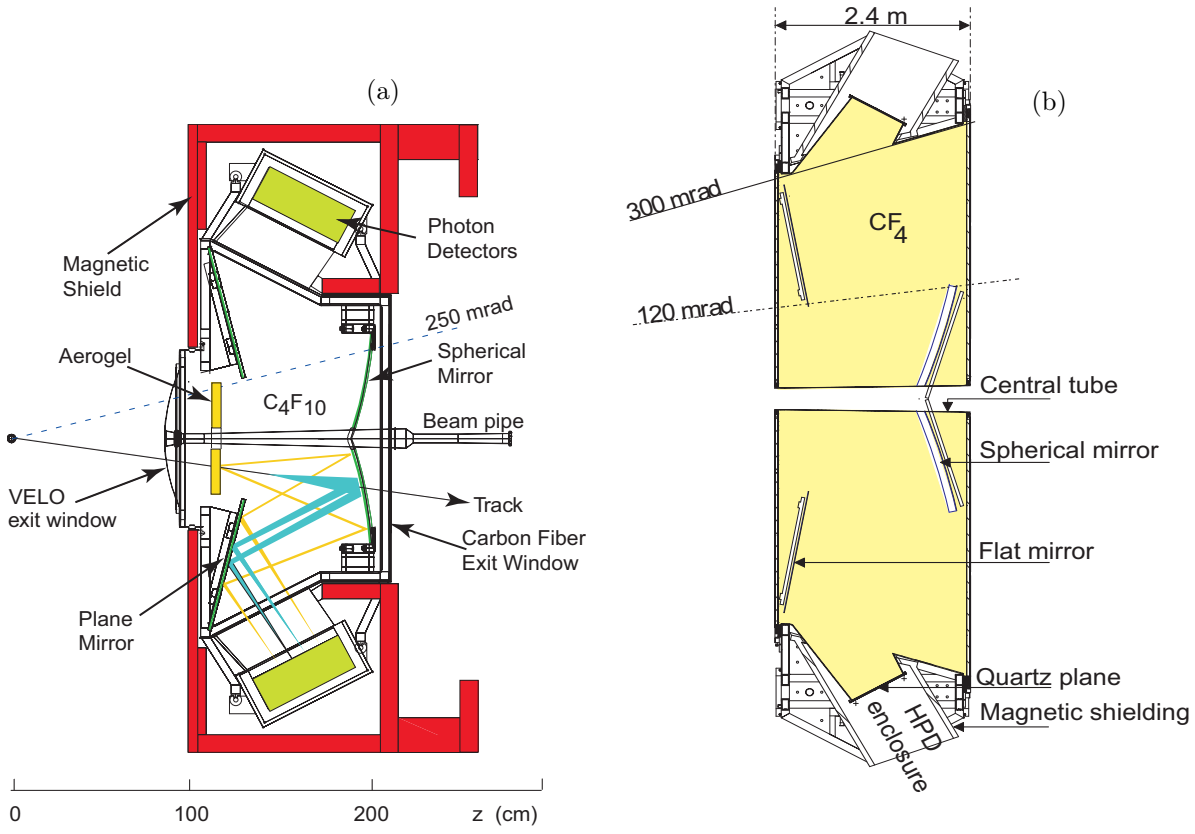


Figure 3.11: (a) The layout of RICH1 (side view); (b) The layout of RICH2 (top view). (from Ref. [55])

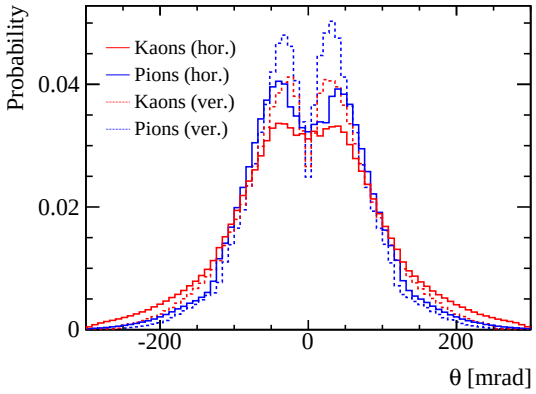


Figure 3.12: Distributions of the horizontal and vertical azimuthal angle θ of the extrapolate position of kaons and pions in RICH2. The tracks have momenta larger than 16 GeV/c and come from $D^0 \to K^+\pi^-$ candidates.

3.2.3 Performance of the RICH Detectors

The spatial resolution of the detected light spots is 0.88 mm (corresponding to 0.35 pixel) in both detectors. The expected Cherenkov angle for different mass hypothesis as a function of the particle momentum as well as the distribution of the measured Cherenkov angles as a function of the measured particle momenta in data are shown in Fig. 3.13. The RICH detectors allow the separation between charged pions and kaons in a momentum range between 5 and 100 GeV/c [44]. The resolution of the Cherenkov angle is 1.618 mrad for C_4F_{10} and 0.68 mrad for CF_4 while it is approximately 6 mrad in Aerogel, but the resolution in Aerogel shows a non-Gaussian behaviour. The average number of photoelectrons emitted by a track is about 5 in Aerogel, between 20 and 25 in C_4F_{10} and about 16 in CF_4 [44].

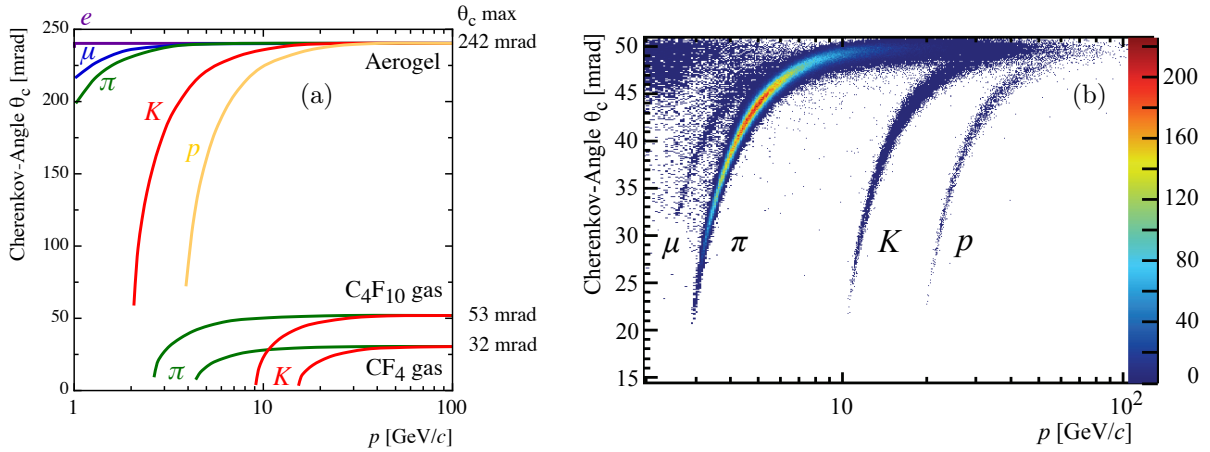


Figure 3.13: (a) Expected Cherenkov angle θ_c in the three radiators (RICH1: Aerogel and C_4F_{10}, RICH2: CF_4) as a function of the particle momentum for different particle types (adapted from Ref. [175]); (b) Reconstructed Cherenkov angle θ_c as a function of the track momentum in data for the C_4F_{10} radiator. The clusters for muons, pions, kaons and protons are well visible. (adapted from Ref. [44])

The information about the particle identification from the RICH detectors is combined with further information from the calorimeters and the Muon System into likelihood classifiers with continuous value ranges comparing two hypothesis about the type of the particle (*e.g.* pion *vs.* electron). The measured particle identification (PID) efficiency based on these classifiers as well as the mis-identification rates for pions, kaons and protons, which are mainly separated through information from the RICH detectors, is shown in Fig. 3.14 as a function of the track momentum. Further details on this topic are discussed in Chapter 7.

3.3 The Calorimeter System

The calorimeter system [55,56] is formed by an electromagnetic calorimeter (ECAL) followed by a hadronic one (HCAL). The system performs several tasks: it delivers information about hadrons, electrons and photons with high transverse momenta for the first trigger stage (*cf.* Sec. 3.5) and is used in the offline analysis to perform energy and position measurements of electrons, hadrons and photons. In addition it also delivers information for PID, especially to distinguish neutral particles like photons and π^0 mesons.
The first stage of the ECAL consists of a scintillator pad detector (SPD) followed by a preshower (PS) detector. The SPD has the aim to distinguish between charged and neutral particles, especially to separate photons and π^0 mesons from electrons while the PS gives a finer longitudinal segmentation in the first stage of the ECAL and helps to separate single photons from photon pairs emitted in a π^0 decay by their longitudinal shower shape.
The combined radiation length of the SPD/PS is $2.5X_0$ while the remainder of the ECAL has a length of $25X_0$ to detect the full energy deposit of the electromagnetic shower. The HCAL has a length of only 5.6 times the hadronic interaction length λ_I due to space limitations. The full energy of the hadronic shower does not have to be detected by the HCAL for trigger purposes.
The transverse segmentation in all calorimeter parts follows the particle density: the SPD/PS and ECAL have cells with a square front face of 40.4 mm in side length in the region closest to the beam pipe, cells of 60.6 mm in side length in an intermediate region and cells of 121.2 mm in side length in the outermost region (*cf.* Fig. 3.15(a)). The HCAL uses cell dimensions of 131.3 mm $\times$ 131.3 mm (inner region) and 262.2 mm $\times$ 262.2 mm (outer region) (*cf.* Fig. 3.15(b)).

3.3.1 The Scintillating Pad/Preshower Detector

The Scintillating Pad/Preshower (SPD/PS) detector consists of a 15 mm thick layer of lead as absorber, sandwiched by two scintillator pad detectors with fine granularity. The first pad detector (SPD) has the aim to detect ionisation deposited by charged particles and so to reject single photons and photon pairs from π^0. Afterward the showering process is initiated in the lead layer. The energy deposition in the second pad detector (PS) gives the finer longitudinal segmentation in the ECAL.
The scintillating material is – as in the ECAL and HCAL – polysterene with a primary and secondary wave-length-shifting dopant[7]. The scintillator pads are wrapped into 0.15 mm thin white, reflecting TYVEK paper to increase the light collection efficiency. The light generated

[7] Type and composition of the wave-length-shifting dopant:
primary: PTP (p-Terphenyl, 1,4-Diphenylbenzene) (SPD/PS: 1.5 %, ECAL: 2.5 %, HCAL: 1.75 %);
secondary: POPOP (1,4-Bis(5-phenyl-2-oxazolyl)benzene) (SPD/PS: 0.04 %, ECAL: 0.01 %, HCAL: 0.05 %)

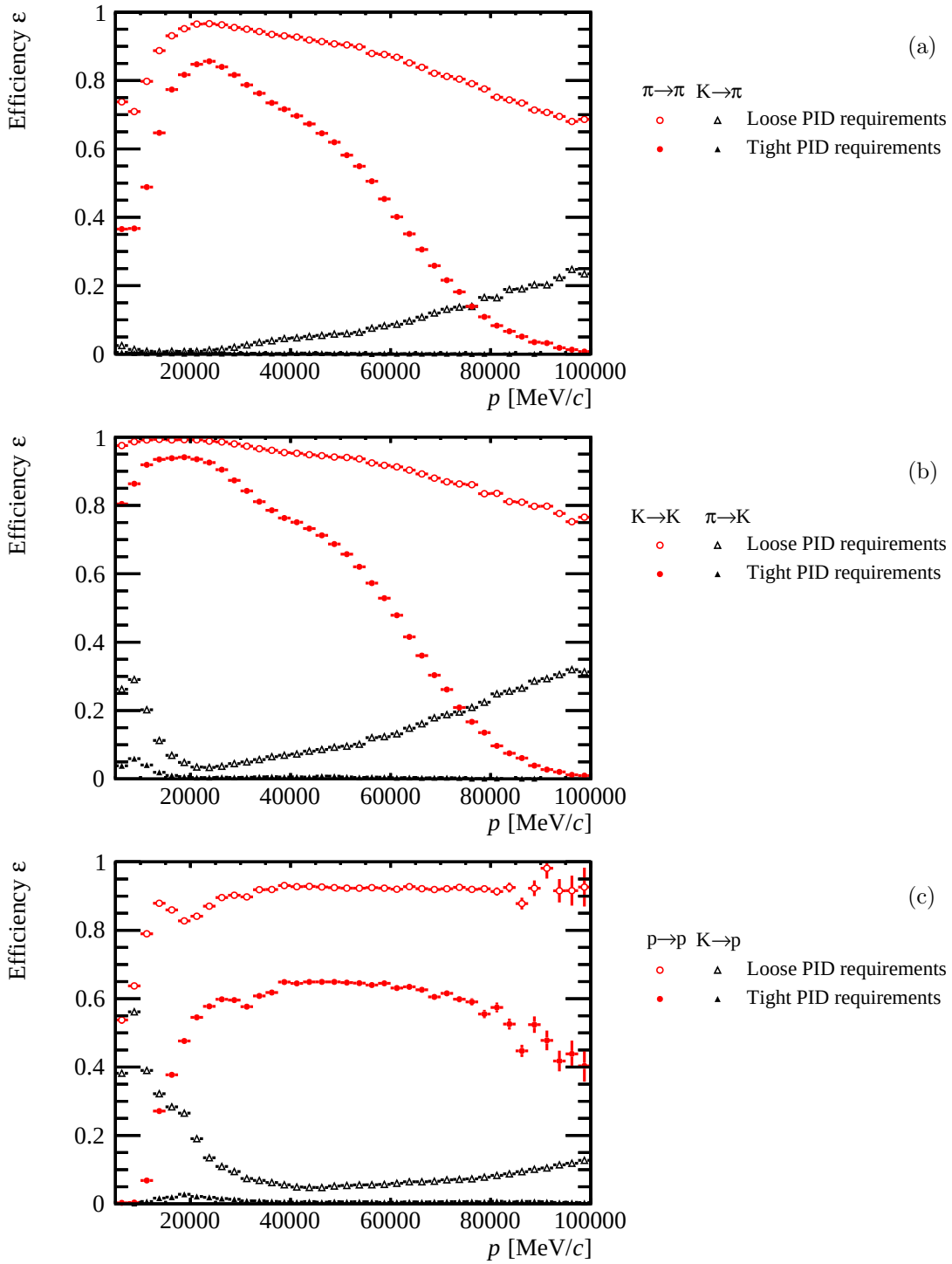


Figure 3.14: Measured PID performance as a function of the particle momentum between 5 and 100 GeV/c using loose and tight PID requirements (for details *cf.* Chapter 7). (a) PID efficiency for pions and the probability for true kaons to be mis-identified as pions; (b) PID efficiency for kaons and the probability for true pions to be mis-identified as kaons; (c) PID efficiency for protons and the probability for true kaons to be mis-identified as protons.

in the scintillating material is collected by wave-length-shifting fibres. Each fibre guides it to a multi-anode photomultiplier tube (MAPMT) detecting the light signal.

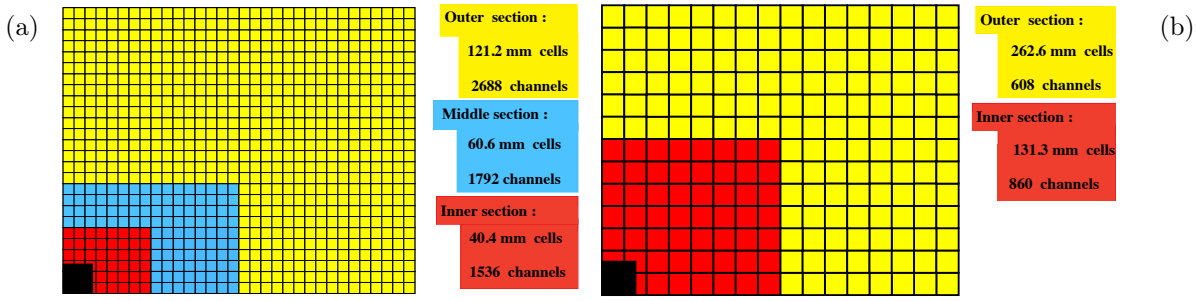


Figure 3.15: Transverse segmentation of (a) the SPD/PS as well as the ECAL and (b) the HCAL. (from Ref. [55])

3.3.2 The Electromagnetic Calorimeter

The ECAL is a sampling scintillator/lead device made of alternating 2 mm thick lead layers as absorber, 0.12 mm thick TYVEK paper sheets and 4 mm thick scintillator tiles. There are 66 lead/TYVEK/scintillator layers stacked in one calorimeter module. The collection and transmission of the scintillator light is done identically as in the SPD/PS, but 64 fibres are collectively read out by a single PMT.
The relative energy resolution of the electromagnetic calorimeter has been measured in an energy range between 10 and 100 GeV as

$$\frac{\sigma_E}{E} = \frac{10\,\%}{\sqrt{E}} \oplus 1\,\% \text{ with } E \text{ in GeV}, \tag{3.2}$$

where the first term is due to the stochastic shower development and the second one due to the instrumental resolution defined among others for example by the detector geometry.

3.3.3 The Hadronic Calorimeter

The HCAL consists of scintillating tiles interleaved with 1 cm thick iron layers oriented parallel to the beam axis. Along this axis the tiles have a length corresponding to the hadronic interaction length in steel (197 mm). They are also wrapped into TYVEK paper to increase the light collection efficiency. The wave-length-shifting fibres are running parallel to the beam axis and transmit the light to the PMTs placed at the downstream end of the HCAL.
The energy resolution described in the same way as for the ECAL has been measured in an energy range between 20 and 100 GeV as

$$\frac{\sigma_E}{E} = \frac{69\,\%}{\sqrt{E}} \oplus 9\,\% \text{ with } E \text{ in GeV}. \tag{3.3}$$

3.3.4 Performance of the Calorimeter System

In terms of PID for electron, the information of the calorimeter system allows to separate electrons from charged hadrons with an efficiency of 92 % giving a mis-identification probability of hadrons

as electrons of 4.5 %. This high performance is achieved by the shower form discrimination based on the SPD/PS.
Further the mass resolution for π^0 reconstructed from two resolved photons in the ECAL is $7.5\,\mathrm{MeV}/c^2$ with an upward shift of the central value with respect to the nominal π^0 mass by $0.3\,\mathrm{MeV}/c^2$.
The separation of single photons from π^0 whose two photons are not resolved (*i.e.* they have overlapping clusters in the ECAL), is about 98 % with a π^0-rejection of about 45 % [33].

3.4 The Muon System

The Muon System [55, 74] is used in the first trigger level to select muons with high transverse momenta p_T and in the later, software based, trigger levels and the offline analysis to identify particles as muons.
There are in total five muon stations (M1-5). The first one (M1) is located between the RICH2 and the calorimeter system. The aim of this station is to improve the p_T resolution for the trigger measurements. To reduce the material upstream of the calorimeter, this station has only half of the detector layers than the other four stations (M2-5).
These are placed downstream from the HCAL. Each of these four stations is followed by a 80 cm thick iron absorber (muon filter) to stop all particles except muons (and neutrinos). Figure 3.16(a) shows the side-view of the Muon System.
The Muon System is based on Multi-wire Proportional Chambers (MWPC). Only the innermost part of M1 uses triple-Gas Electron Multipliers (GEM) detectors to cope with the high occupancy and irradiation in this region.
Both detector types use an Argon/CO_2/CF_4 gas mixture. The cathodes of both detectors and the anodes of the triple-GEMs are formed by pads. The MWPCs use 30 µm thick gold-plated tungsten anode wires. The signal is read-out on both anodes and cathodes.
The anode wires or anode pads run along the y-axis (x-position sensitive) and the cathode pads along the x-axis (y-position sensitive), which allows an x-y-sensitivity by performing a logical AND operation on the anodes and cathodes in a single detector.
The stations M2-5 (M1) consist each of four (two) single detector layers, which are connected in the read-out by a logical OR operation to provide redundancy and increase the detection efficiency.
Similar to the calorimeter system, the Muon System has a finer granularity closer to the beam pipe than in the outer regions. There are four regions, R1-4, different in granularity, as shown in Fig. 3.16(b). The granularity decreases in all stations following the decreasing particle density towards the outer regions of the detector. M2 and M3 have the same granularity, which is twice as high as in M1 while also M4 and M5 have the same granularity, but which is half as high as in M1.

3.4.1 Performance of the Muon System

The identification of muons is based on a likelihood classifier as mentioned in Sec. 3.2.3. In addition there is a binary classifier called *IsMuon* to select muon candidates. This classifier is based on the number of hits in the muon stations M2-5 associated to a track as charged non-muon particle should not penetrate through the calorimeter and the several muon filters into all the stations. Table 3.2 shows the requirements on hits in the muon stations associated to a

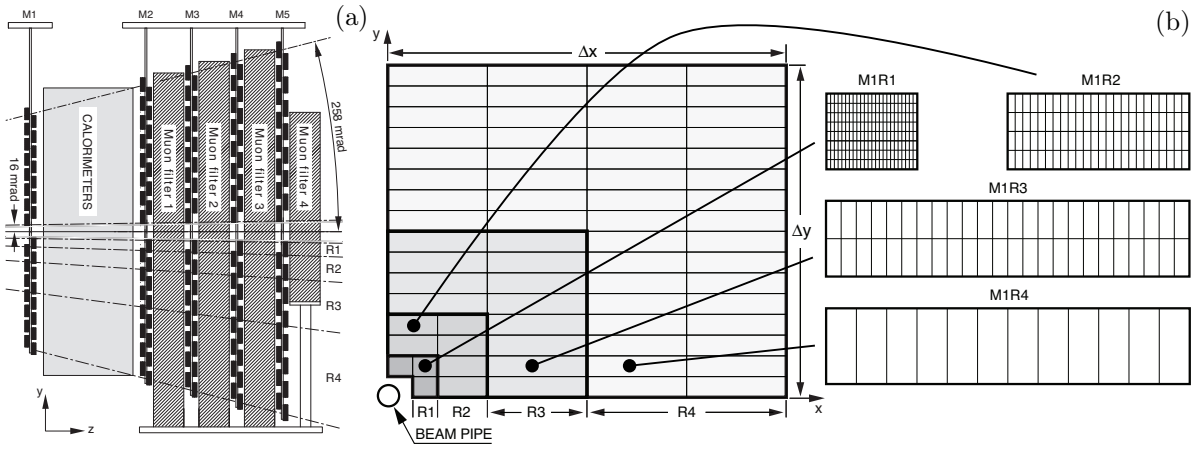


Figure 3.16: (a) The layout of the five muon stations and the muon filters forming the Muon System; (b) Front view of a quadrant of a muon station (left). Each rectangle represents one chamber and the different shading shows the different regions R1-4. The right part of (b) shows for M1 the division of each chamber type into logical pads in the four regions R1-4. The lines show where which type of division is used. In the case of M2/M3 (M4/M5) the number of pad columns for each chamber is double (half) the number in the corresponding type of M1. (from Ref. [55])

track. As the penetration power of a muon is momentum-dependent, the requirements are also momentum-dependent [65, 186].

Table 3.2: Requirements of the binary muon classifier on hits associated to a track. [65, 186]

Momentum range	**Requirements on the muon hits**[a]
$3\,\text{GeV}/c < p < 6\,\text{GeV}/c$	M2 and M3
$6\,\text{GeV}/c < p < 10\,\text{GeV}/c$	M2 and M3 and (M4 or M5)
$10\,\text{GeV}/c < p$	M2 and M3 and M4 and M5

[a] Showing from which muon station a hit must be associated to the track.

The performance of this classifier (efficiency and mis-identification probability) is shown as a function of the track momentum in Fig. 3.17(a). It identifies very efficiently (around 97 %) muons while rejecting background very successful (around 99 %), especially for track momenta larger than $20\,\text{GeV}/c$.
The performance of the identification based on the likelihood classifier separating muons and pions is depicted in Fig. 3.17(b).

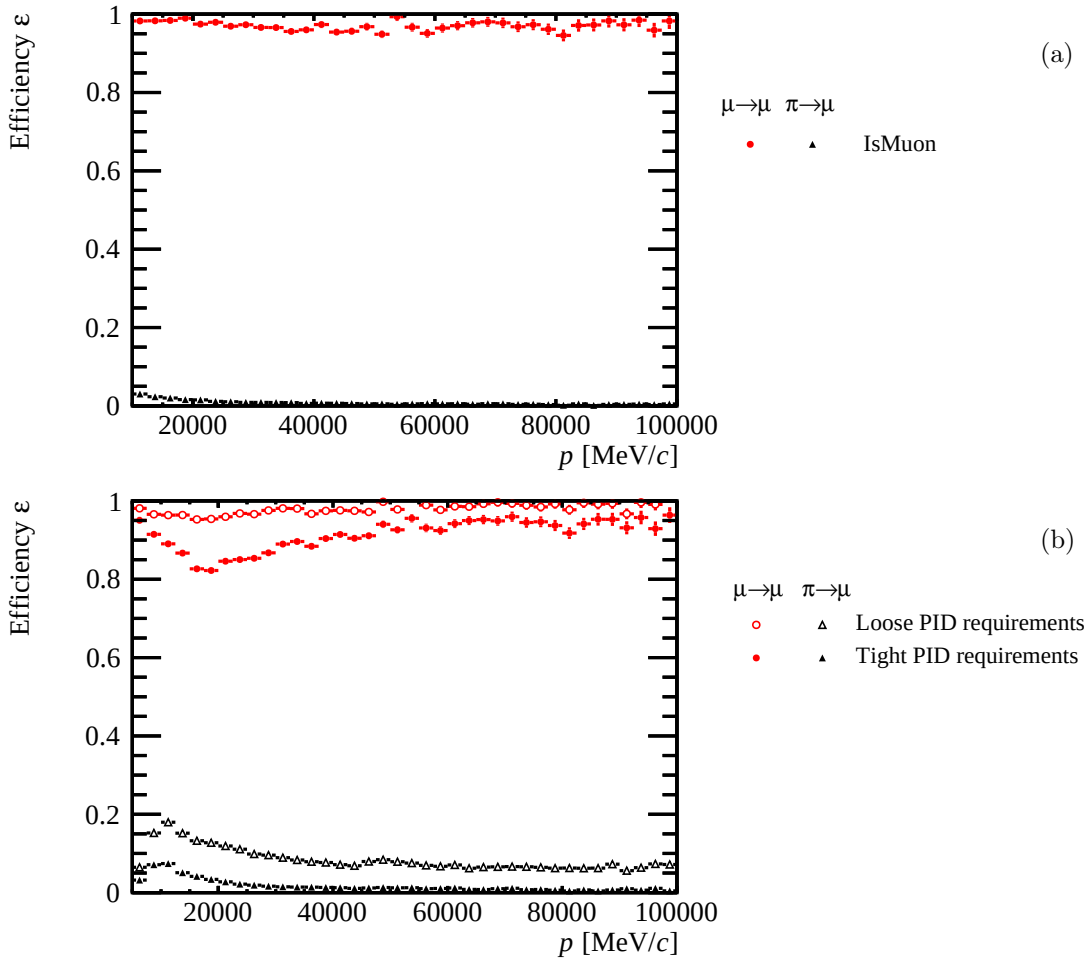


Figure 3.17: (a) Efficiency of the binary IsMuon classifier and its mis-id probability for pions as a function of the particle momentum; (b) Efficiency of the likelihood-classifier to separate muons and pions and its mis-id probability for pions as a function of the particle momentum.

3.5 The Trigger System

The bunch crossing frequency at the interaction point of LHCb is up to 40 MHz. Due to the bunch structure and the lower luminosity than in the general purpose detectors at LHC, the frequency of pp interactions visible to the LHCb detector is 10 MHz [55].
The goal of the trigger system [55, 62] is to reduce it to 2 kHz. At this rate the data from the sub-detectors are stored for offline analysis. The trigger system of LHCb consists of three trigger stages. The first stage (Level-0 trigger; L0) is a hardware-based trigger while the other two levels (High-Level Trigger 1/2; HLT1/2) are software-implemented.
The L0-trigger uses information from the pile-up detector, integrated in the VELO, the calorimeters and the Muon System. The pile-up detector was designed to reject events with multiple pp interactions, but under the conditions during the data taking from 2010 to 2013, LHCb also recorded bunch crossings with several primary vertices as the experience showed that such events can be used for physics analyses. Hence the detector is used to give a first measurement of the position of the primary vertices and to detect collisions of a particle in one of the beams with residual gas molecules in the beam vacuum (beam-gas collisions).
The information from the calorimeters and the Muon System is used to trigger on hadrons, photons, electrons and muons with high transverse energy E_{T} or momentum p_{T}. The decision of the L0-trigger is done using customised electronics based on FPGAs. Its latency, *i.e.* the time elapsed between the interaction and the arrival of the L0-trigger decision at the front-end electronics of the sub-detectors is 4 µs, where half of it is used to process the data in the L0-trigger to derive the decision and the other half is used by the delays in the front-end electronics, the cable delays and the time-of-flight of the particles. The L0-trigger has to reduce the event rate in the data processing to 1 MHz, at which rate the remainder of the detector is read out. During the data taking periods in 2011 and 2012 the maximum output rate from the L0-trigger has been 1.1 MHz.
The High-Level Trigger uses an Event Filter Farm (EFF) consisting of about 29'000 logical computing cores to refine the trigger decision from the L0-trigger [29]. As the 1 MHz output rate of the L0-trigger is still too high for a full reconstruction of the event, a first stage of the HLT (HLT1) performs a partial reconstruction of the event by reconstructing the segments in the VELO and the tracking stations and matching them to the calorimeter clusters and Muon System segments detected in the L0-trigger. This procedure allows to confirm and refine the trigger decision from the L0. The output rate of the HLT1 is about 50 kHz and hence low enough to perform a full reconstruction of the event, which is then used in the second HLT stage (HLT2) [157, 158].
In the HLT2, the event reconstruction is done as close as possible to the offline reconstruction [36]. This stage uses to select events by a set of inclusive and exclusive algorithms, which are tailored to select specific decays or set of decays such as decays of b and c hadrons to two, three or four charged final state particles [233] or detached $J/\psi(1S)$ candidates as decay products from b hadrons [36]. In total about half of the trigger band width is used for inclusive hadronic triggers while another 25 % are taken from triggers on leptons and the remaining quarter is used by exclusive triggers, mainly for decays of charm hadrons [29].

3.6 The LHCb Software

There are several relevant taylor-made software packages for analyses at LHCb [123]. As some will play an important role in the later chapters, a short overview of those and their purposes is given here:

Package	Description
Boole	Simulation of the detector response for simulated events.
Brunel	Reconstruction of events from data or Monte Carlo simulation based on the information from the sub-detectors.
DaVinci	Analysis framework to perform selections and calculations on reconstructed events.
Gaudi	Core software package for the data processing in LHCb
Gauss	Generation of simulated data samples [78,118] with different event generators and tools to simulate particle decays as well as simulation of the transition of the particles through the detector using the Geant4 toolkit [48,52]. For standard MC simulation Pythia6 [216] or Pythia8 [217] are used as event generators while the decay of particles is simulated with EvtGen [187] and final state radiation (FSR) with Photos [2].
Moore	High-Level-Trigger application to run the trigger algorithms for HLT1 and HLT2 as well as to emulate the trigger decisions on simulated data.
Panoramix	Event display for data and simulation.
Vetra	Offline analysis and monitoring of the VELO and the ST.

II

Radiation Damage in the LHCb Silicon Tracker

The essential question is
how much harm has the age done to a person's character.

Mark Twain (1835 - 1910)

4

Silicon Strip Detectors and Radiation Damage

A crucial point in the operation of silicon sensors at hadron colliders is a proper understanding of the impact of radiation on the sensor properties, the monitoring of these effects and a reliable prediction for the future evolution of the irradiated sensors.
The monitoring and the prediction of the future radiation damage can help to take necessary measures in terms of operation to ensure a good long-term performance of the sensors.
This chapter has the goal to give an introduction into the basic properties of silicon micro-strip sensors (Sec. 4.1), the impact of radiation damage (Sec. 4.2) and the tools to predict the development of the effect from radiation damage (Sec. 4.3).

4.1 The Basic Principle of Silicon Sensors

4.1.1 The Design of Silicon Sensors

The fundamental principle of silicon sensors is that ionising particles traversing silicon excite along their flight path electrons from the valance to the conductor band, creating electron-hole pairs inside the material. If these charge carriers are moved by an applied external voltage to the surface of the silicon, they induce an electric signal on the read-out electrodes, which can be read out with low-noise electronics.
The number of electron-hole pairs created by a minimal-ionising particle in 1 mm of silicon is about 70'000 while the number of thermally created electron-hole pairs at room temperature in 1 mm^3 of silicon is about three orders larger. Therefore the signal is completely drowned by statistical fluctuations, unless the thermally created charge carriers can be removed from the silicon bulk.
The solution is the design of the silicon sensor as a p-n junction, corresponding to the layout of a diode: an interface of p-doped and n-doped silicon is created. The surplus of highly mobile electrons as majority charge carriers in the n-doped region and the surplus of holes in the p-doped region lead to a diffusion of electrons from the n-doped to the p-doped region as well as of holes

in the opposite direction followed by a recombination of electrons and holes. This mechanism creates at the interface a charge-carrier free zone as well as an electric field from the n-doped to the p-doped region counteracting the diffusion. In the equilibrium, the force of the electric field and the diffusion compensate each other and a stable charge-carrier free zone (*depletion zone*) is created.
The electric field gradient across the p-n junction with the spatial coordinate, x', running from the p- to the n-doped region and $x' = 0$ at the interface, is in the equilibrium approximately

$$\begin{aligned} \frac{\mathrm{d}E}{\mathrm{d}x'} &= -\frac{\mathrm{d}^2\Phi}{\mathrm{d}x'^2} = \frac{\rho(x')}{\varepsilon\varepsilon_0} \\ &= \begin{cases} \dfrac{n_A q}{\varepsilon\varepsilon_0} & \text{for } -d_p < x' < 0 \\ -\dfrac{n_D q}{\varepsilon\varepsilon_0} & \text{for } 0 < x' < d_n \\ 0 & \text{else,} \end{cases} \end{aligned} \tag{4.1}$$

where n_D (n_A) is the density of donor (acceptor) atoms on the n-side (p-side) and d_n and d_p are the thicknesses of the depletion zones on the two sides. The variable q is the electron charge and $\varepsilon\varepsilon_0$ is the effective permittivity of silicon.
The intrinsic potential difference, V_{intr}, between the two sides of the interface can be calculated from Eq. (4.1) with the proper boundary conditions of $E(-d_p) = E(d_n) = 0$ and is

$$V_{\mathrm{intr}} = \frac{q}{2\varepsilon\varepsilon_0}(n_D d_n^2 + n_A d_p^2). \tag{4.2}$$

As the whole sensor must be electrically neutral, there is also the condition

$$n_A d_p = n_D d_n. \tag{4.3}$$

For the following discussion, the effective doping concentration, n_{eff}, is defined as

$$n_{\mathrm{eff}} = \frac{n_D d_n^2 + n_A d_p^2}{(d_n + d_p)^2}. \tag{4.4}$$

In the case of the ST where the sensors are built from an n-doped bulk with p^+-doped implants (*cf.* Fig. 4.1), the concentration values for donor atoms are of the order of $\mathcal{O}(10^{12})$ cm^{-3} and for acceptor atoms of the order of $\mathcal{O}(10^{15})$ cm^{-3}, *i.e.* $n_A \gg n_D$ and therefore with Eq. (4.3) $d_n \gg d_p$. Therefore, n_{eff} can be approximated by $n_{\mathrm{eff}} \approx n_D$.

4.1.2 Sensor Depletion

If an external reversed-bias voltage V_{bias} is applied, the thickness of the depletion zone will increase to

$$d = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \cdot (V_{\mathrm{bias}} + V_{\mathrm{intr}}) \cdot \left(\frac{1}{n_D} + \frac{1}{n_A}\right)} \tag{4.5}$$

$$\approx \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \cdot \left(\frac{V_{\mathrm{bias}} + V_{\mathrm{intr}}}{n_{\mathrm{eff}}}\right)}. \tag{4.6}$$

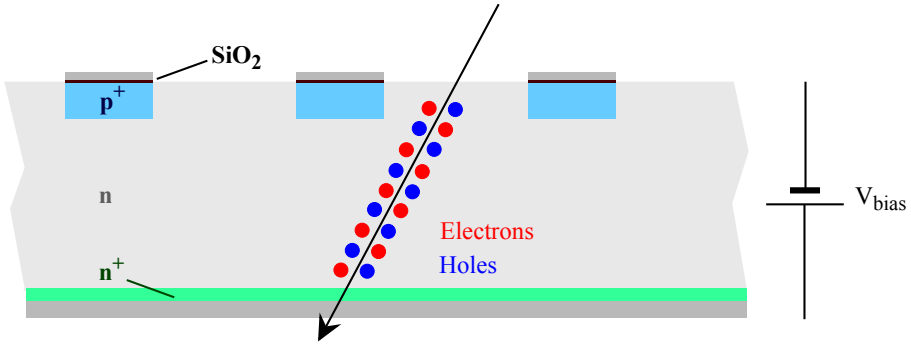


Figure 4.1: Schematic of a silicon sensor of p-on-n type as used in the ST. The coordinates x and x' are described in the text.

When the depletion zone extends over the full thickness, D, of the silicon sensor, the sensor is called *fully depleted* and the corresponding voltage is called *depletion voltage* V_{depl} with

$$V_{\text{depl}} = \frac{q}{2\varepsilon\varepsilon_0} n_{\text{eff}} D^2 - V_{\text{intr}}. \tag{4.7}$$

Typically, V_{depl} is much larger than V_{intr}, which can thus be neglected.
The electric field generated in the silicon sensor by this applied revesed-bias voltage is

$$E(x) = \begin{cases} \dfrac{2V_{\text{depl}}}{D}\left(1 - \dfrac{x}{D}\right) + \dfrac{V_{\text{bias}} - V_{\text{depl}}}{D} & \text{for } x < d \\ 0 & \text{else,} \end{cases} \tag{4.8}$$

where x is the coordinate across the sensor thickness with $x = 0$ at the side of the sensor connected to the higher potential of the bias voltage (*i.e.* in normal p-on-n type sensors the backplane). This electric field causes the drift of the charge carriers inducing the electric signal on the read-out electrodes. The velocity v of a charge carrier at point x is given by

$$v(x) = \mu \cdot E(x), \tag{4.9}$$

where the mobility, μ, it is about 14 m^2/Vs for electrons and 4.5 m^2/Vs for holes at room temperature [221]. The mobility shows a temperature dependence of $T^{-2.4}$ for electrons and $T^{-2.2}$ for holes leading to mobilities of 16 m^2/Vs for electrons and 5.2 m^2/Vs for holes at the operational temperature of the sensors in the ST of 8°C [227]. The doping concentration only affects the mobility in highly doped sensors, *i.e.* $n_i = \mathcal{O}(10^{18})$. The charge collection time $t(x)$ for charge carriers produced at a position x is thus

$$t(x) = \int_x^0 \frac{\mathrm{d}x'}{v(x')}. \tag{4.10}$$

Silicon sensors of p-on-n type are operated with $V_{\text{bias}} > V_{\text{depl}}$ (overdepletion) to guarantee an electric field across the whole sensor. Typical drift times for sensors operated as in the ST are of the order of $\mathcal{O}(10\text{ ns})$.

The depletion of the sensor also changes its capacitance C. It can be approximated by describing the sensor as a parallel-plate condensator with capacitance

$$C = \begin{cases} \varepsilon\varepsilon_0 \dfrac{A}{d} \approx A\sqrt{\dfrac{\varepsilon\varepsilon_0 q \cdot n_{\text{eff}}}{2V_{\text{bias}}}} & \text{for } V_{\text{bias}} < V_{\text{depl}} \\ \varepsilon\varepsilon_0 \dfrac{A}{D} & \text{for } V_{\text{bias}} \geq V_{\text{depl}}. \end{cases} \tag{4.11}$$

Therefore the measured value of $1/C^2$ as a function V_{bias} can be used to determine V_{depl} where $1/C^2$ starts to be flat. A larger sensor capacitance corresponds also to a larger electronic noise of the front-end amplifiers (*capacitive noise*). Thus, the noise decreases due to the sensor depletion. But due to additional contributions to the noise from the resistance of the signal path in the read-out electronics, called *thermal noise* and due to statistical fluctuations in the leakage current (*cf.* Sec. 4.1.3), called *shot noise*, it is not possible to obtain a measurement of C via the observed noise in the sensor.

4.1.3 The Leakage Current

The direct current due to thermally created electron-hole pairs in a diode operated in the reverse direction is called *leakage current* (or reverse current). On the one hand effects in context with interfaces or edges of the sensor can contribute to the leakage currents. On the other hand impurities and defects can also increase the leakage current by creating additional energy levels between the valence and the conductor band. The increase of the leakage current caused by radiation damage thus consists of two components: there is the *bulk current* caused by the additional energy levels increasing the number of thermally excited electron-hole pairs. Further, there are *surface currents* due to radiation induced effects at the silicon-SiO_2 interface (*cf.* Sec. 3.1.3).
The sensor geometry and the radiation type that the ST sensors are exposed to suggest that the leakage currents in the ST are dominated by the bulk current I. As it is mainly caused due to the excitation of electron-hole pairs, it also shows a temperature dependence, which can be described by

$$\frac{I(T_1)}{I(T_2)} = \left(\frac{T_1}{T_2}\right)^2 \exp\left[\frac{E_g}{2k_B}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right], \tag{4.12}$$

where T_1 and T_2 are two arbitrary temperatures and E_g is the band gap in silicon [221].

4.2 Radiation Damage Effects in Silicon Sensors

4.2.1 Radiation Damage Types

As pointed out above there are two types of radiation damage in silicon sensors. One type, called *surface damage* is due to X- or gamma-rays as well as charged particles ionising silicon atoms or producing electron-hole pairs. As long as they are produced in the bulk, the charge carriers are collected by the electrodes. If they however are produced in the insulating oxide layer, they cannot be collected by the electrodes and are trapped in the oxide. This increases the positive charge of the oxide. Therefore the amount of surface damage depends on the dose the sensor is exposed to.

The second type, called *bulk damage* is in general caused due to changes in the silicon lattice induced by the radiation. Particles in the radiation cause non-ionising energy loss (NIEL) interactions with a recoil atom in the lattice. This can lead to the displacement of the atom creating an interstitial and a vacancy in the lattice, called a *Frekel pair*. The interstitial as well as the vacancy can diffuse through the lattice. Either the two recombine, causing beneficial annealing or each of them forms a permanent defect with other imperfections in the lattice. The damage efficiency of a particle does not only depend on its type, but also on its energy. Therefore damage coefficients have been measured for various particle types over a wide range of energies [219, 220]. For the sake of comparison, the flux of particles is normalised to a reference type and energy, which are usually slow neutrons with a kinetic energy of 1 MeV. This normalised fluence, $\Phi_{1\text{ MeV-n,eq}}$, is called the *1-MeV neutron equivalent fluence* and will be later used in the simulation of radiation damage in the ST.
Especially particles with small energies (*i.e.* shortly before they are captured) produce Frekel pairs at a high density, creating clusters of them. A cluster of interstitials and vacancies can more easily lead to a permanent defect as the probability of recombination of the Frekel pair is smaller [198, 199].

4.2.2 Change of Sensor Properties

There are three main properties of a silicon sensors that change due to radiation damage: the leakage current increases as described above, n_{eff} and thus V_{depl} change, and the total charge collection efficiency (for a fully depleted sensor) decreases due to trapping of charge carriers at the induced defects in the lattice.
In most of the cases, the change of V_{depl} limits the lifetime of the detector as at a certain point the required bias voltage exceeds the break-through voltage of the diode. But a carefully considered design can extend the lifetime of the sensor. Designs of the strip geometry avoiding sharp ages causing high electric fields or the usage of guard rings allow to apply high bias voltages. A further measure to reduce the radiation damage effects is to keep the sensors cold.
As in the case of the ST the doping can be chosen such that the irradiation first decreases the effective doping until the so-called point of *type inversion* is reached and afterwards the effective doping and thus the depletion voltage raises. For detectors in a very high radiation environment, n^+-on-n type sensors can be used, which can be operated partially depleted after type inversion.

4.3 Modeling of Radiation Damage Effects

Detailed studies of irradiated silicon sensors led to phenomenological models allowing to estimate the evolution of the sensor performance in a known radiation environment and thus to optimise the operation parameters [143, 150, 198, 234]. But usually, the radiation environment is not known to a good precision requiring a monitoring of the radiation damage effects to compare the measured evolution with the expected one. A detailed description of the models can be found in Ref. [198].
In the following, models are described for the effects on the leakage current as well as on the depletion voltage. These models include the direct impact of radiation on the considered quantities as well as the effects of annealing processes of the induced defects.

4.3.1 Effects on the Leakage Current

If the leakage current is dominated by the bulk current, its change, ΔI_{leak}, is proportional to the total received fluence Φ as the number of induced intermediate energy levels is proportional to the number of induced defects. Further ΔI_{leak} is also proportional to the volume, V, of the irradiated silicon as the leakage current is proportional to the number of thermally excited electron-hole pairs.
The proportionality is described by the current related damage rate, α, leading to the formula

$$\Delta I_{\text{leak}} = \alpha \cdot \Phi \cdot V. \tag{4.13}$$

This linear behaviour describes measurements well for a fluence range in terms of $\Phi_{1\text{ MeV-n,eq}}$ between 10^{11} and 10^{15} cm^{-2} [198].
The impact of annealing can be described by introducing a time- and temperature-dependence of the proportionality factor α. For short time scales after the irradiation (minutes to days), the time dependence of α can be described by the sum of decreasing exponential functions and of a constant. The latter describes the effect of stable damage. The time scales τ_i of the exponential functions are temperature dependent [197]. The temperature dependence can be described by the Arrhenius relation, representing the temperature dependence of the reaction time, τ, in simple chemical reactions, as

$$\frac{1}{\tau} = k_0 \exp\left(-\frac{E_a}{k_B T_a}\right), \tag{4.14}$$

where k_0 is called the frequency factor, and E_a is the activation energy of the reaction [66].
To describe the annealing behaviour over longer time scales of a month to a few years, the proportionality factor in the model can be reduced to the sum of a single exponential function and a constant term [196]

$$\alpha(t, T_a) = \alpha_{I,0} + \alpha_{I,1} \exp\left(-\frac{t}{\tau_{I,1}(T_a)}\right). \tag{4.15}$$

This simplified model is used in the next chapter to compare the leakage currents measured in the ST during LHC Run I with predictions. The choice of the parameters used to describe the measurements will be also discussed in the next chapter.

4.3.2 Effects on the Depletion Voltage

The effects of radiation damage on the depletion voltage can be described by the so-called *Hamburg model*, introduced in Ref. [198]. The Hamburg model describes the change of the effective doping concentration, Δn_{eff}, as a function of the fluence, normalised to $\Phi_{1\text{ MeV-n,eq}}$, taking into account annealing processes depending on the time, t, since the irradiation and the temperature T_a during this period.
Three different effects lead to a change in the effective doping concentration: the first contribution, n_c comes from the stable damage to the silicon and depends purely on the fluence. The second contribution, n_a, describes the annealing of the defects and depends besides $\Phi_{1\text{ MeV-n,eq}}$ also on t and T_a while the third contribution, n_r, incorporates the reverse (or anti-) annealing and depends also on t and T_a. So Δn_{eff} is given by

$$\Delta n_{\text{eff}} = n_c(\Phi_{1\text{ MeV-n,eq}}) + n_a(\Phi_{1\text{ MeV-n,eq}}, t, T_a) + n_r(\Phi_{1\text{ MeV-n,eq}}, t, T_a). \tag{4.16}$$

In the following, the three components are briefly described.

Stable Damage Term

The stable damage term, n_c, is associated to the effective (incomplete) removal of donor atoms and the addition of stable acceptor atoms due to a change in the band structure by the radiation. The removal of donor atoms can be described by a change in the effective doping concentration proportional to $[1 - \exp(-c\Phi_{1\text{ MeV-n,eq}})]$ as the amount of donor atoms is decreasing due to the irradiation while the addition of stable acceptor atoms is proportional to $\Phi_{1\text{ MeV-n,eq}}$. Therefore, the stable damage part can be described by

$$n_c(\Phi_{1\text{ MeV-n,eq}}) = n_{c,0}[1 - \exp(-c\Phi_{1\text{ MeV-n,eq}})] + g_c\Phi_{1\text{ MeV-n,eq}}, \tag{4.17}$$

where the parameters $n_{c,0}$, c and g_c must be determined from irradiation campaigns of silicon sensors.

Annealing Term

In a time period directly after their irradiation, silicon sensors show a change in n_{eff} that goes in the opposite direction to the impact of the irradiation itself. This behaviour can be interpreted as an annealing of the induced defects in the silicon (*e.g.* by recombination).
The corresponding contribution to the change of the effective doping, n_a, can be described as an exponential decrease

$$n_a(\Phi_{1\text{ MeV-n,eq}}, t, T_a) = \Phi_{1\text{ MeV-n,eq}}\, g_a \exp\left(-\frac{t}{\tau_a(T_a)}\right), \tag{4.18}$$

where t is the time since the irradiation with the fluence $\Phi_{1\text{ MeV-n,eq}}$, and g_a the average introduction rate, determined from irradiation campaigns. The variable τ_a defines the time scale of the process and is temperature dependent. This dependence can also be described by the Arrhenius relation

$$\frac{1}{\tau_a} = k_{a,0} \exp\left(-\frac{E_{aa}}{k_B T_a}\right), \tag{4.19}$$

with the parameter E_{aa} corresponding to the activation energy of the process and $k_{a,0}$ to the frequency factor. This relation makes the process strongly temperature dependent: for a temperature of -10° C τ_a is about a year while for a temperature of 80° C the time scale decreases to a few minutes.
The analytic description in Eq. (4.18) can obviously only be applied if the irradiation and the annealing happen sequentially, otherwise a differential approach is required.

Reverse Annealing Term

The last contribution comes from the combination of individual defects as mention above leading to a reverse annealing and a change in the effective doping. This contribution can be parametrised by

$$n_r(\Phi_{1\text{ MeV-n,eq}}, t, T_a) = \Phi_{1\text{ MeV-n,eq}}\, g_r \left(1 - \frac{1}{1 + t/\tau_r}\right), \tag{4.20}$$

where t is again the time since the irradiation with the fluence $\Phi_{1\text{ MeV-n,eq}}$, and g_r the average introduction rate, determined from irradiation campaigns.
The time scale of this process is given by the parameter τ_r, which also has a temperature dependence that can be described by the Arrhenius relation (*cf.* Eq. (4.14)), with activation energy E_{ar} and frequency factor $k_{r,0}$. The reverse annealing has a larger frequency factor than the annealing causing a longer time scale (500 years at -10° C, 1.5 hours at 80° C).
Thus, high temperature in a short period after the irradiation can cure radiation damages while silicon sensors need to be kept cold over longer periods to suppress the diffusion of defects allowing them to build formations with other defects in the lattice.

As mentioned above, these analytical formulae can only be applied in a sequential cycle of irradiation and annealing. During operation of the LHC this is not the case as irradiation extends over long periods of time. Therefore, a differential approach must be applied, which can be described by

$$\begin{aligned}\left.\frac{\mathrm{d}n_{\text{eff}}}{\mathrm{d}t}\right|_{t=0} &= n_{c,0}c\exp(-c\Phi_{1\text{ MeV-n,eq}})\frac{\mathrm{d}\Phi_{1\text{ MeV-n,eq}}}{\mathrm{d}t} \\ &\quad +\Phi_{1\text{ MeV-n,eq}}\left(\frac{g_r}{\tau_r(T_a)}-\frac{g_a}{\tau_a(T_a)}\right)+\frac{\mathrm{d}\Phi_{1\text{ MeV-n,eq}}}{\mathrm{d}t}(g_a+g_c).\end{aligned} \tag{4.21}$$

This approach is used in the next chapter within a 4th order Runge-Kutta algorithm to predict Δn_{eff} and thus the evolution in the depletion voltage. The choice of the parameters to describe the data is also discussed in the next chapter.

5

In-situ Measurements of Radiation Damage in the LHCb Silicon Tracker

The monitoring of the radiation damage in the ST is based on two methods: the first one uses the observed change in the leakage current while the second method is based on measurements of the depletion voltage using dedicated Charge Collection Efficiency (CCE) scans. The two methods are described in Secs 5.2 and 5.3, respectively. The first section of the chapter describes the fluence simulation used for the radiation damage predictions while the final section discusses the results of the measurements.

5.1 Fluence Simulation

The different fluences (neutrons, charged hadrons, electrons and positrons as well as the 1-MeV neutron equivalent fluence) and the dose in the LHCb cavern have been simulated using the Fluka simulation code [77, 146]. The fluences and dose as a function of x and y have been evaluated at different z-positions. The simulation has been tuned based on the measured dose at different locations in the cavern during Run I of the LHC. Separate simulations for the two centre-of-mass energies $\sqrt{s} = 7$ and 8 TeV have been performed.

5.1.1 Fluence in the Tracker Turicensis

The most important simulation for these studies is that of the 1-MeV neutron equivalent fluence. It is calculated for the TT with a scoring of 1 cm in the x- and y-direction while the scoring in the z-direction is 1 mm. For all detection layers of TT, a single z-position is used, which is located in the centre of the TT station at $z = 254.8$ cm and covers the range $(x, y) \in ([-60, +60]\text{ cm} \times [-60, +60]\text{ cm})$. Figure 5.1 shows the two-dimensional 1-MeV neutron equivalent fluence maps for pp-collisions at $\sqrt{s} = 7$ and 8 TeV as well as their one-dimensional profiles for $x = 0$ cm and $y = 0$ cm. There is a difference of about three orders of magnitude between the highest and

the lowest flux in the region covered by the TT sensors. The highest 1-MeV neutron equivalent fluence in the TT and thus also in ST is about 10^{12} cm^{-2} per 1 fb^{-1}. The corresponding value for the VELO of LHCb is about a factor 10 higher [46], while the highest fluence in the region covered by silicon sensors in ATLAS and CMS is approximately a factor 50 higher [17].
The expected highest fluence in the ST up to a possible upgrade of LHCb after 2019 is of the order of 10^{13} cm^{-2}.

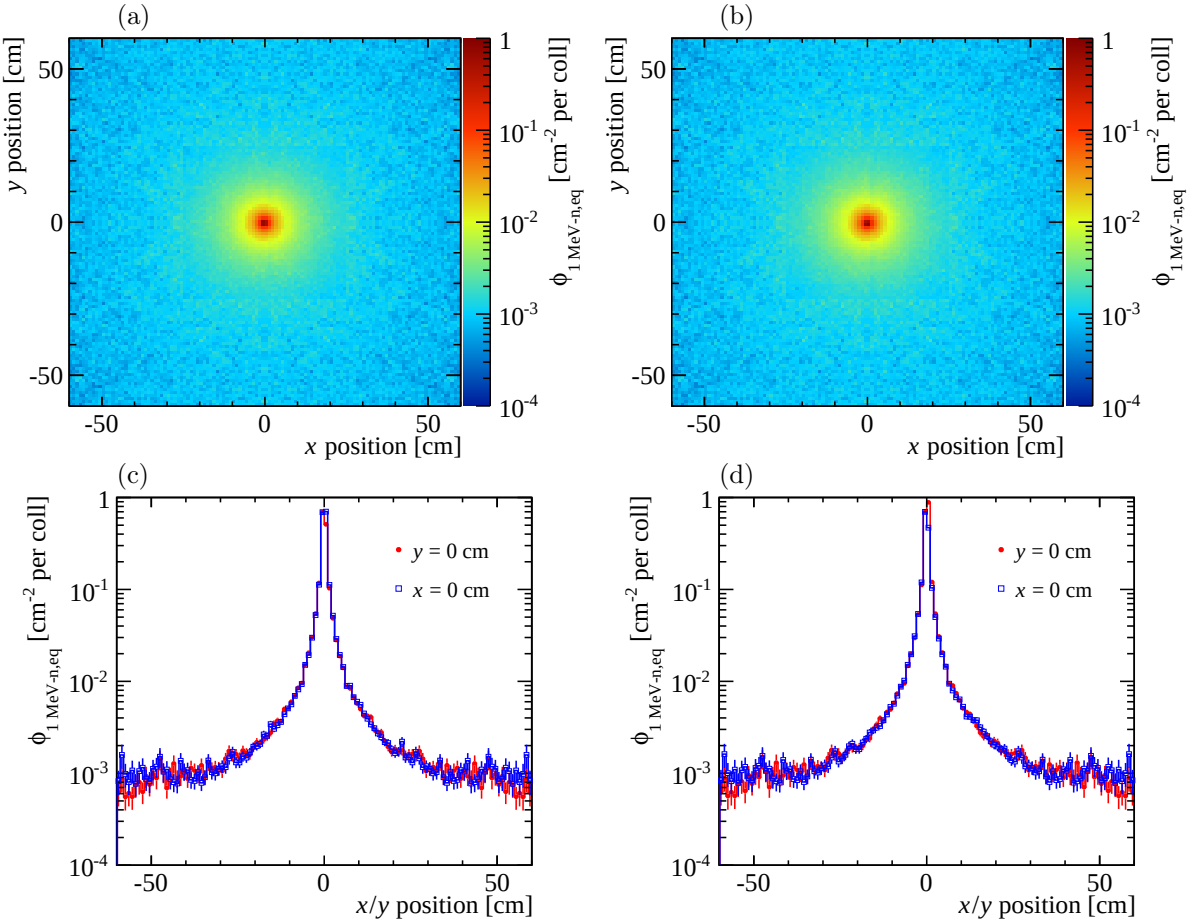


Figure 5.1: Simulated 1 MeV-neutron equivalent fluence in the TT per pp collision at (a) $\sqrt{s} = 7$ TeV and (b) $\sqrt{s} = 8$ TeV. The profile of the fluence maps from FLUKA for $x = 0$ cm and $y = 0$ cm are shown in (c) for $\sqrt{s} = 7$ TeV and (d) for $\sqrt{s} = 8$ TeV.

5.1.2 Fluence in the Inner Tracker

For IT, the simulation is performed for statistical reasons using a coarser scoring in x and y of 5 cm. Furthermore, the used simulation only covers the IT detector boxes in x and y. The z-position of the projection is the centre of the IT station T3 at $z = 904.0$ cm as the radiation damage studies in the IT mainly concentrate on this station. The obtained two-dimensional 1-MeV neutron equivalent fluence maps for pp-collisions at 7 and 8 TeV are shown in Fig. 5.2. The highest fluence value in the IT is about one order of magnitude smaller than in the TT.

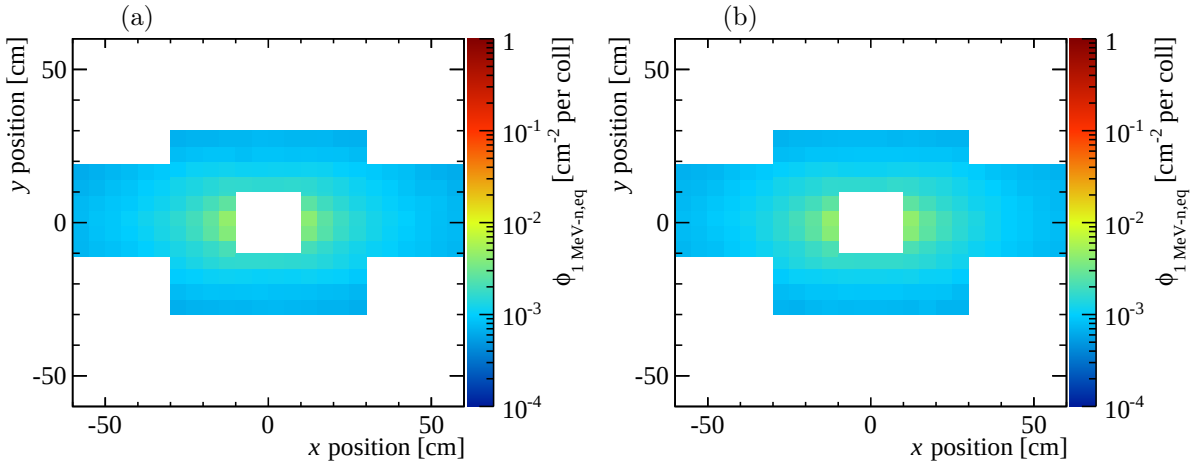


Figure 5.2: 1 MeV-neutron equivalent fluence simulated in FLUKA in the IT station T3 per *pp* collision at (a) $\sqrt{s} = 7\,\mathrm{TeV}$ and (b) $\sqrt{s} = 8\,\mathrm{TeV}$.

For both sub-detectors, the relative statistical uncertainty $\sigma_{\phi_{1\ \mathrm{MeV\text{-}n,eq}}}/\phi_{1\ \mathrm{MeV\text{-}n,eq}}$ on the simulation as a function of the distance r from the beam axis can be parametrised by a power-law function of the form

$$\sigma_{\phi_{1\ \mathrm{MeV\text{-}n,eq}}}/\phi_{1\ \mathrm{MeV\text{-}n,eq}} = a_0 \cdot (r/r_0)^{\alpha} \tag{5.1}$$

with $a_1 = 0.7 \times 10^{-2}\ \mathrm{cm}^{-2}$, $r_0 = 1\ \mathrm{cm}$ and $\alpha = 0.918$ (*cf.* Fig. 5.3).

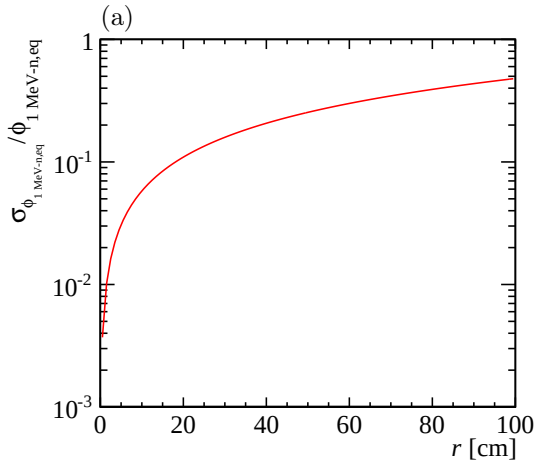


Figure 5.3: Relative uncertainty on the FLUKA simulation as function of the distance r from the beam axis described by a power-law function.

5.2 Leakage Current Measurements

The most direct measurement to monitor the radiation damage in a silicon sensor in usage is the measurement of its leakage currents. In the ST, the leakage currents are defined as the maximal currents observed during an LHC fill in the different bias voltage channels. A bias voltage

channel in the IT is connected to four read-out sectors (4 or 8 sensors) of the same type and in the same station and region (*i.e.* A-/C-side, top, bottom), but not necessarily in the same layer. The bias voltage channels in the TT are connected to either 1, 2, 3, 4, 9 or 12 sensors. The number of sensors is increasing with the distance from the beam axis. This allows to tune the bias voltage with a higher granularity in the innermost part of the detector, which is affected the most by radiation damage. Table 5.1 lists the different possibilities of active silicon volume connected to a single bias voltage channel in the TT and IT.

Table 5.1: Summary of the HV partitoning in the TT and IT.

Sub-detector	**Geometry (single sensor)**	**Number of sensors**	**Volume [cm^3]**
IT	11.00 cm × 7.60 cm × 0.032 cm	4	10.70
	11.00 cm × 7.60 cm × 0.041 cm	8	27.42
TT	9.64 cm × 9.44 cm × 0.050 cm	1	4.55
		2	9.10
		3	13.65
		4	18.20
		9	40.95
		12	54.60

The obtained current data are cleared of erraneous measurements by removing zero-current measurements and measurements exceeding the previous measurement by a factor of three or more. Also fills with special configurations (*e.g.* higher bias voltage) are removed. Furthermore, only the inner two TT layers U and V are studied as the layers TTaX and TTbX showed very high currents during the first part of LHC Run I. Figures 5.4 and 5.5 show the leakage currents per silicon volume, I_{leak}/V, as a function of time and as a function of the integrated delivered luminosity, respectively. The currents have been normalised to a temperature of $T_a = 8°$ C, based on Eq. (4.12) with $E_{\text{g}} = 1.21$ eV.

The measured ambient temperatures are shown as a function of time in Fig. 5.6. In the case of the TT the values are obtained from two temperature sensors placed in the centre of the detector box. They are directly mounted on the support rails of the half-modules in the layers TTaU and TTbV above the beam pipe. The average temperature of the two sensors is taken as ambient temperature value. In the IT, the temperature inside each of the twelve detector boxes is measured individually. The precision of the ambient temperature measurements is assumed to be 2° C.

Figures 5.4 and 5.5 also show the predicted change in I_{leak}/V. This prediction is based on the model described in Sec. 4.3.1, applied on the simulated 1 MeV-neutron equivalent fluence described in Sec. 5.1, the actual running conditions (*e.g.* integrated luminosity, time of fills, fill length) and the temperature measurements. The parameters used in this simulation are listed in Tab. 5.2.

In the observed leakage current, time periods of short technical stops (about one week every three months) or longer shutdowns (in the winter season) are visible. These time periods lead to decreases in I_{leak}, which correspond to annealing processes in the silicon. The predicted evolution

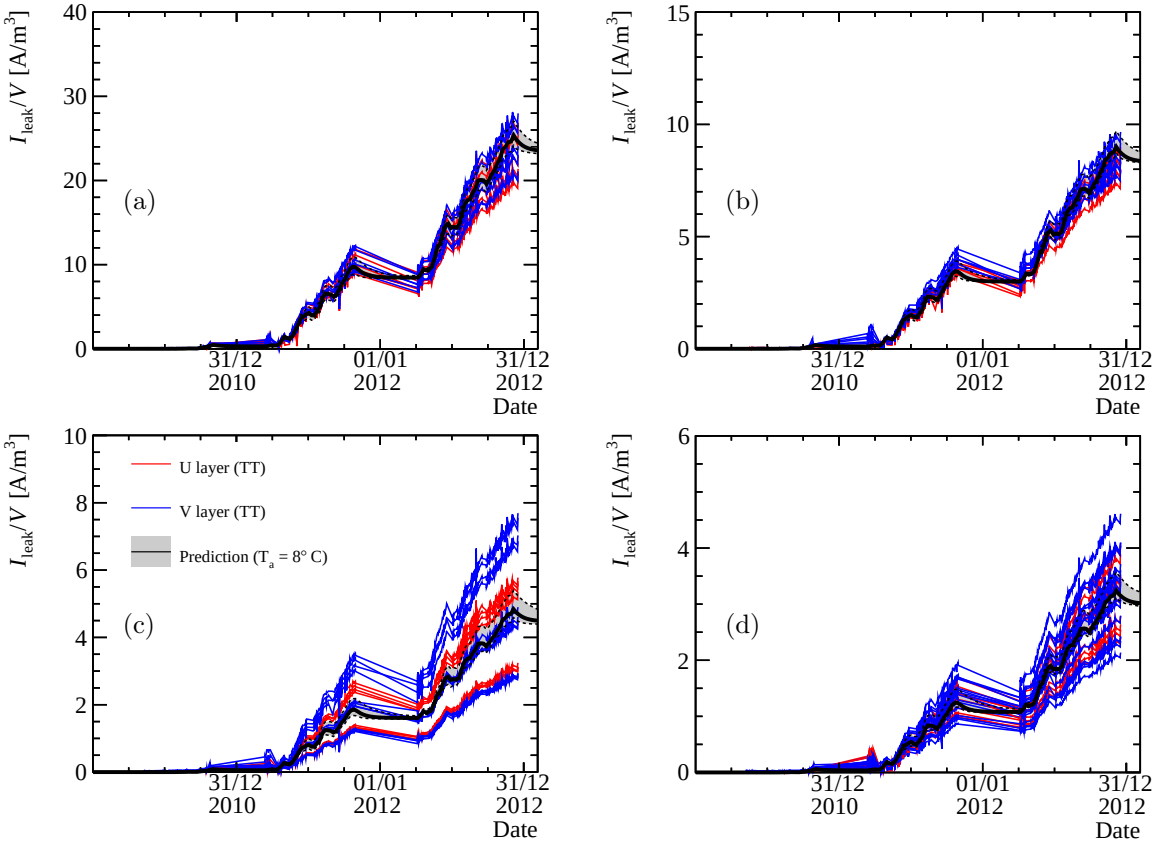


Figure 5.4: Leakage currents per silicon volume, I_{leak}/V, in the TT layers U (red lines) and V (blue lines) as a function of time. HV channels are connected to (a) one-sensor, (b) two-sensor, (c) three-sensor and (d) four-sensor read-out sectors. The black line shows the prediction based on the average 1-MeV neutron equivalent fluence in the sensors obtained from FLUKA simulations. The grey band and the dotted lines show the uncertainty on these predictions based on the uncertainty on the model parameters, on the FLUKA simulation and on the temperature measurements. It does not take into account the variation in the fluence across the compared sensors. Periods of annealing during winter shutdowns and technical stops are visible in the data and the predictions.

Table 5.2: Parameters for the leakage current simulation.

Parameter	**Value**
$\alpha_{I,0}$	$(6.67 \pm 0.09) \times 10^{-17}$ A/cm
$\alpha_{I,1}$	$(7.23 \pm 0.06) \times 10^{-17}$ A/cm
k_0	$(4.2 \pm 0.5) \times 10^{13}$ s^{-1}
E_a	(1.11 ± 0.05) eV

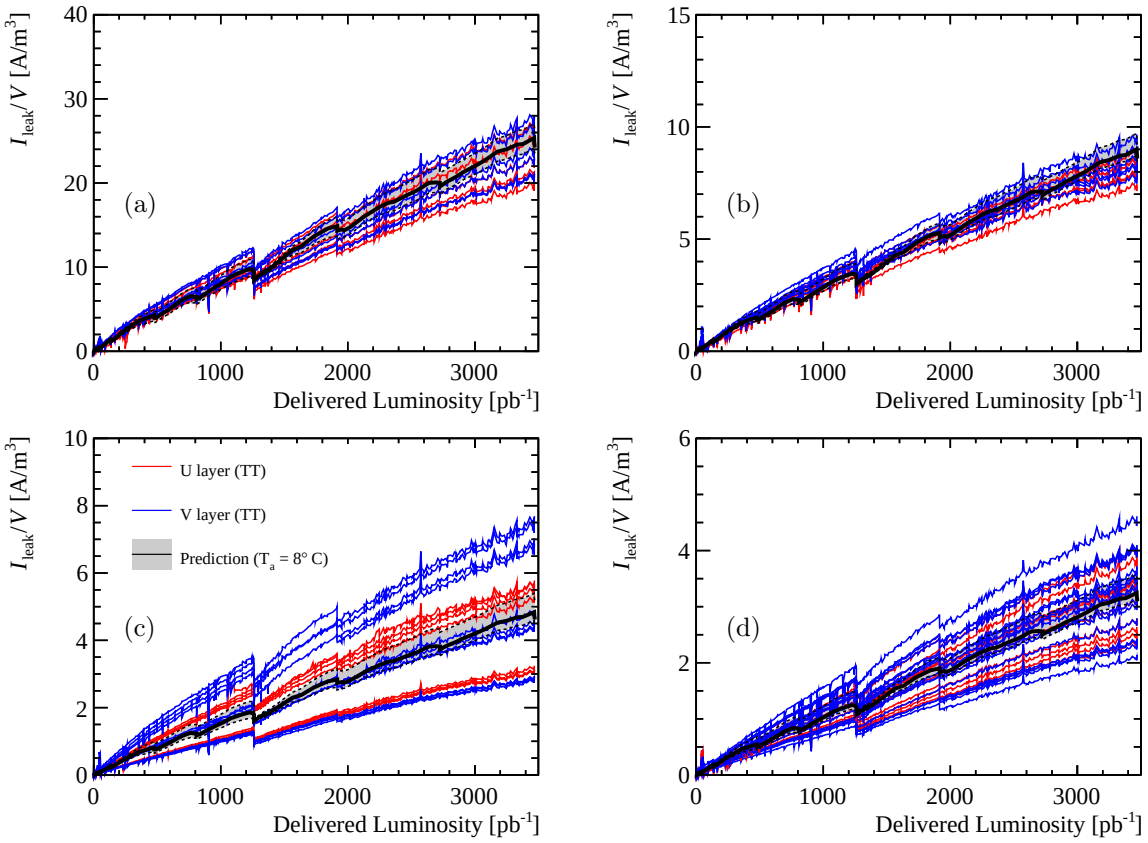


Figure 5.5: Leakage currents per silicon volume, I_{leak}/V, in the TT layers U (red lines) and V (blue lines) as a function of the integrated delivered luminosity. HV channels are connected to (a) one-sensor, (b) two-sensor, (c) three-sensor and (d) four-sensor read-out sectors. The black line shows the prediction based on the average 1-MeV neutron equivalent fluence in the sensors obtained from FLUKA simulations. The grey band and the dotted lines show the uncertainty on these predictions based on the uncertainty on the model parameters, on the FLUKA simulation and on the temperature measurements. It does not take into account the variation in the fluence across the compared sensors. Periods of annealing during winter shutdowns and technical stops are visible in the data and the predictions.

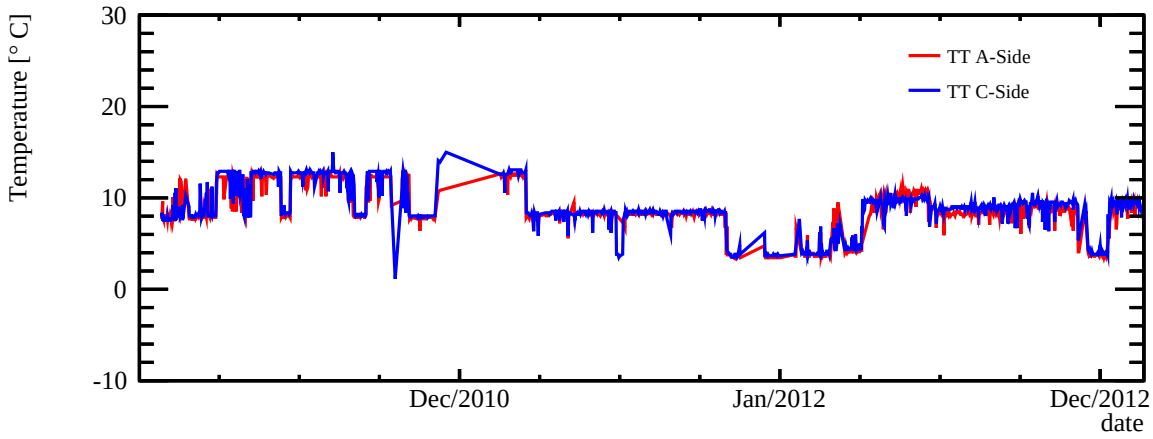


Figure 5.6: Ambient temperature in the TT detector box as a function of time. The measurements are from two temperature sensors placed on the support rails of the top center half-modules of the layers TTaU and TTbV very close to the beam pipe. One is placed in the A-side region of the TT, the other in the C-side region. Fewer temperature measurements have been taken over the shutdown periods in the winter months.

of I_{leak}/V is in good agreement with the measured one.
The measured values of I_{leak}/V as a function of time and integrated luminosity in the IT are shown in Figs 5.7 and 5.8, respectively. The values of I_{leak} are again normalised to $T_a = 8°$ C. Figure 5.9 shows the ambient temperature in the IT detector boxes as a function of time. The comparison with the predicted change in I_{leak}/V shows also in this sub-detector a good agreement.

5.3 Depletion Voltage Measurements

As discussed in the previous chapter, radiation damage to silicon sensors changes their effective doping, which leads to a change in the depletion voltage, V_{depl}, as described by Eq. (4.7). An in-situ measurement of V_{depl} via the sensor capacitance or the sensor noise is due to technical reasons not possible.
A more suitable approach for in-situ monitoring of the effective doping concentration is to measure the amount of charge induced by ionising particles that is collected by the silicon sensor as a function of V_{bias}. In first order this amount should saturate for $V_{\text{bias}} > V_{\text{depl}}$ while it is increasing with V_{bias} for values below V_{depl}.
The idea of the applied approach to measure V_{depl} is to use dedicated collision data that have been recorded in so-called Charge Collection Efficiency (CCE) scans. Tracks of charged particles reconstructed in these data allow to distinguish signal from noise hits in the tracking stations, which is essential as the average signal height is very small for a low bias voltage. In this way also biases from the clustering algorithms are avoided.
For each applied bias voltage the most probable value of the signal height distribution is measured for different sampling times. This allows to obtain for each bias voltage the signal pulse shape and thus to get a measure representing the amount of collected charge produced by the traversing particle. The depletion voltage can be extracted from the behaviour of the amount of

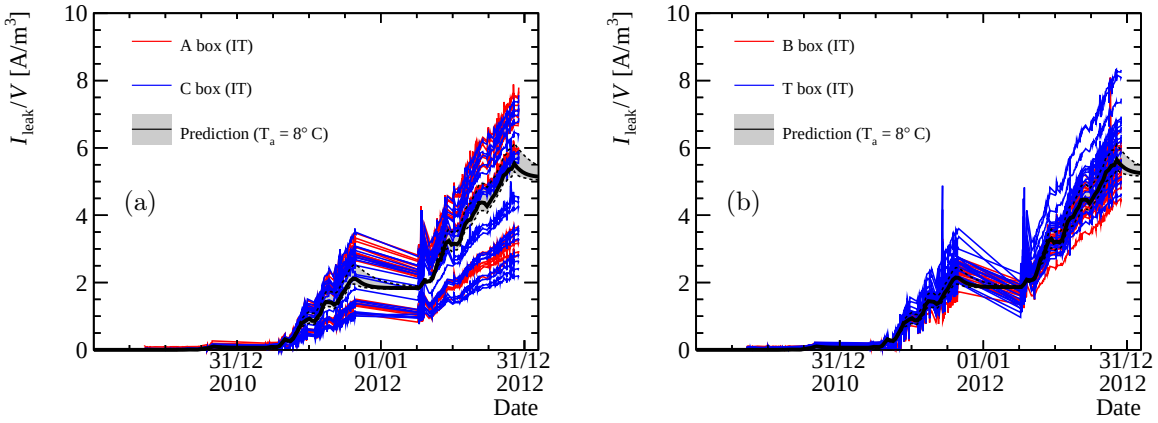


Figure 5.7: Leakage currents per silicon volume, I_{leak}/V, in the IT boxes (a) A-side (red lines) and C-side (blue lines) and (b) bottom (red lines) and top (blue lines) as a function of time. The black line shows the prediction based on the average 1-MeV neutron equivalent fluence in the sensors obtained from FLUKA simulations. The grey band and the dotted lines show the uncertainty on these predictions based on the uncertainty on the model parameters, on the FLUKA simulation and on the temperature measurements. It does not take into account the variation in the fluence across the compared sensors. Periods of annealing during winter shutdowns and technical stops are visible in the data and the predictions.

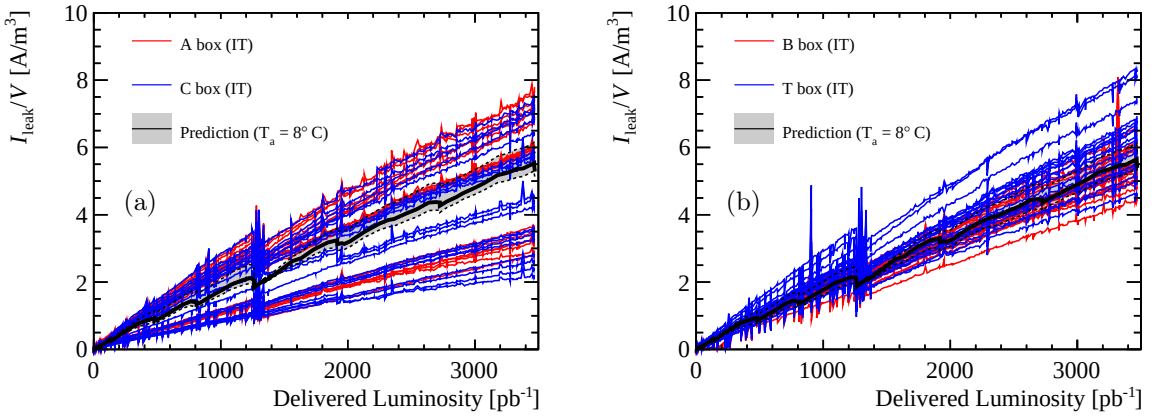


Figure 5.8: Leakage currents per silicon volume, I_{leak}/V, in the IT boxes (a) A-side (red lines) and C-side (blue lines) and (b) bottom (red lines) and top (blue lines) as a function of the integrated luminosity. The black line shows the prediction based on the average 1-MeV neutron equivalent fluence in the sensors obtained from FLUKA simulations. The grey band and the dotted lines show the uncertainty on these predictions based on the uncertainty on the model parameters, on the FLUKA simulation and on the temperature measurements. It does not take into account the variation in the fluence across the compared sensors. Periods of annealing during winter shutdowns and technical stops are visible in the data and the predictions.

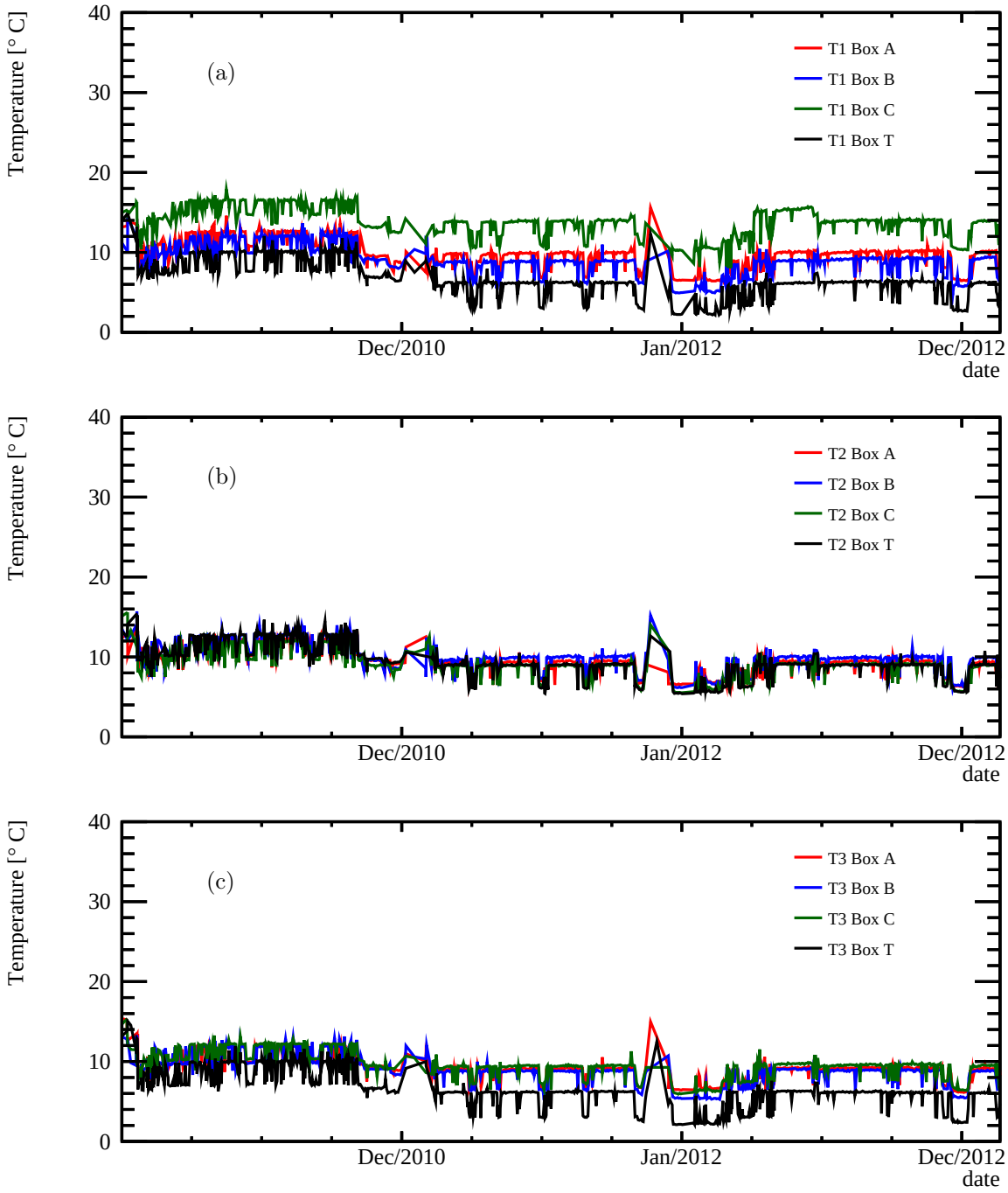


Figure 5.9: Ambient temperature in the IT detector boxes (A: A-side, B: bottom, C: C-side, T: top) shown for (a) T1, (b) T2, (c) T3. The temperature in each detector box is measured by an individual sensor. Fewer temperature measurements have been taken over the shutdown periods in the winter months.

collected charge as a function of the applied bias voltage.

5.3.1 Charge Collection Efficiency Scans

Charge Collection Efficiency (CCE) scans taking place three to four times a year (typically before or after a technical stop of the LHC). Table 5.3 summarises the CCE scans performed so far.

Table 5.3: Summary of CCE scans in 2011, 2012 and 2013.

Date	**Fill Nr.**	$\sqrt{s}$	**Integrated Luminosity at $\sqrt{s}$ = 7 and 8 TeV [fb^{-1}]**	**Magnet Polarity**
2011-03-14	1616	7 TeV	0.04+0.00	Down
2011-07-14	1944	7 TeV	0.48+0.00	Up
2011-09-07	2083	7 TeV	0.81+0.00	Up
2011-10-25	2252	7 TeV	1.24+0.00	Up
2012-04-05	2472	8 TeV	1.26+0.01	Down
2012-07-02	2797	8 TeV	1.26+0.66	Up
2012-09-28	3108[a]	8 TeV	1.26+1.46	Up
2013-01-22	3478[b]	5 TeV	1.26+2.21	Down

[a] Part of this scan was also done with Fill 3109.

[b] This is a proton-lead fill with the proton beam propagating from the VELO towards the muon stations.

In these scans, V_{bias} of one layer in the TT (TTaU) and one layer in the IT (T3X2) is varied. The V_{bias} steps in the TT are [60,100,125,150,175,200,225,250,275,300,400] V while for the IT they are [20,60,90,105,120,135,150,175,200,250,300] V. To be able to perform a measurement of the signal pulse shape for the different values of V_{bias}, a timing scan for each voltage step is performed where the signal sampling time of the Beetle chip (*cf.* Sec. 3.1.3) is shifted by $\delta t = -6.25, -3.13, 0.00, 3.13, 6.25, 9.37$ ns with respect to its nominal value. All other ST layers are kept at their normal bias voltage and sampling time during these scans[1].
This means that the other layers can be used to perform tracking as shown in Fig. 5.10(a). The reconstructed trajectories of the charged particles allow to estimate the position of their hits in the scanned ST layers. Typically, during the CCE scans in the ST also the VELO performs a CCE scan in every fifth sensor while all other VELO stations are kept at the normal configuration [46]. The OT does not perform any special measurements during these scans and is in its nominal configuration.
The data taken in these scans are read out without zero-suppression, *i.e.* also the ADC values of strips not assigned to a cluster (*i.e.* signals on adjacent strips caused by the a traversing particle) are read out.

The raw data is processed offline using dedicated algorithms from the VETRA software package

[1] Actually the high voltage partition in the IT implies that the bias voltage is simultaneously changed in all four layers of the station T3.

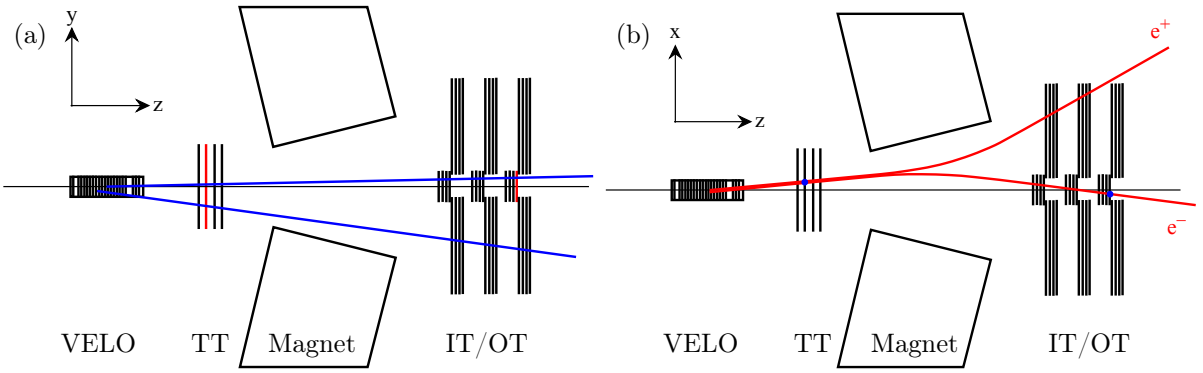


Figure 5.10: (a) Scheme of the LHCb tracking system showing the idea of CCE scans: the black detector layers in the tracking stations and four out of every five of the VELO stations are used in the tracking while the bias voltage of the red layers in the TT and IT is scanned. The reconstructed tracks are extrapolated to the scanned layers (red), pointing to the expected hit position in that layer. (b) Photons converted to electron-positron pairs in the VELO material lead to a two times larger charge signal in the TT while they are separated in the T stations due to the bending by the dipole magnet.

(Vetra v13r2 with Rec v14r3 and Brunel v44r1) to obtain an unbiased distribution of signal height for each read-out sector in TTaU and T3X2. In a first step, the normal track reconstruction is performed after removing all clusters in TTaU and T3X2. Figure 5.11 shows the distribution of the track fit quality and ghost probability[2] for those tracks.
The reconstructed long tracks[3] are then extrapolated to the two scanned layers using a parabolic track extrapolator [87]. Figure 5.12 shows the distribution of the extrapolated track positions in the TT (TTaU) and IT (T3X2). While the distribution across the area covered by sensors is almost uniform in the IT, there is a difference of more than three orders of magnitude between the innermost region and the outer part in the TT.
Using the extrapolated track position in the scanned detector layers, an estimate of the signal amplitude in the scanned layer is obtained by summing the ADC values of the read-out strip closest to the extrapolated point and the neighbouring strip on each side (TT) or the three neighbouring strips on each side (IT). This difference between the TT and the IT is chosen as tracks traverse the T stations typically with a larger angle to the beam axis than TT. This means that the charge carriers are produced in a sensor volume with a larger x-projection in the IT than in the TT. Furthermore, the track extrapolation in the IT is less precise as the scanned layer is the most downstream one of the tracking system while TTaU is preceded by the layer TTaX.

5.3.2 Track Quality Requirements

Figure 5.13 shows the signal height distributions in a TTaU read-out sector for $V_{\text{bias}} = 60$ V and for $V_{\text{bias}} = 400$ V, both for the nominal Beetle sampling time ($\delta t = 0$ ns). For high bias voltage values there is a clear separation between the signal height distribution from real hits, giving the

[2] The ghost probability is a multivariate classifier with values between 0 and 1 to distinguish real tracks from ghost tracks (*i.e.* combination of unrelated hits or track segments in the tracking stations) using kinematical variables as well as the difference between observed and expected hits in the tracking sub-detectors as input. The classifier is built in such a manner that its distribution is flat for ghost tracks.

[3] The tracking algorithm uses in the T1-3 also clusters coming from the OT.

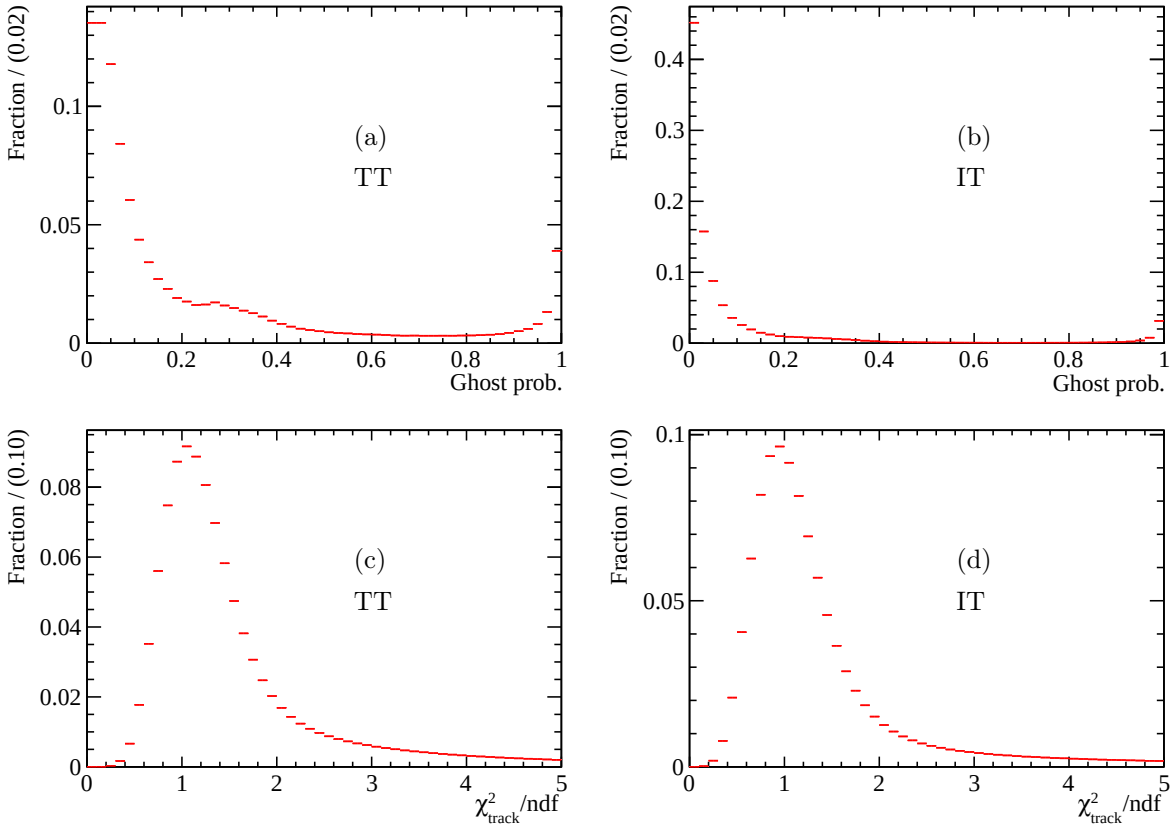


Figure 5.11: The ghost probability distributions of all tracks reconstructed in the CCE scans (a) for the TT and (b) for the IT as well as the track χ^2 distributions of these tracks (c) for the TT and (d) for the IT. The distributions are different between the TT and the IT as the TT sample contains also tracks reconstructed with hits in the OT.

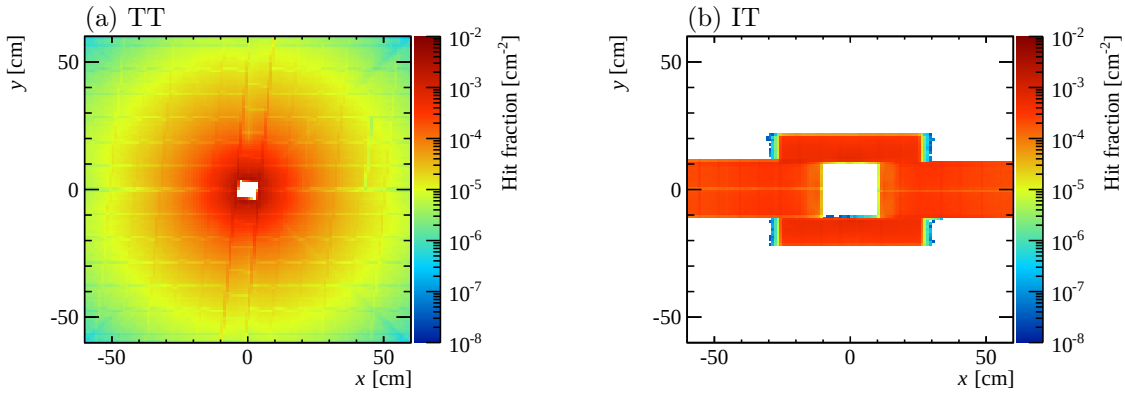


Figure 5.12: Distribution of extrapolated track positions after applying the quality requirement (a) in the TT and (b) in the IT. While the distribution in the IT is almost uniform there is a difference of more than three orders of magnitude between the innermost part of the TT and the outer region, which reflects the particle density in the LHCb detector. In both sub-detectors, the observed distributions show reflections of the structure of the sensors, localised inefficiencies in the detectors and the effect of surrounding objects (*e.g.* diagonal lines in the TT due to the beam pipe suspensions).

expected shape of a Landau distribution, and a signal height distribution from noise centered around zero due to ghost tracks and imperfections in the track extrapolation. In the case of low bias voltage values, the two distributions overlap, causing a possible bias on the most probable values of the signal distribution extracted using a maximum-likelihood fit method. Requirements on the track fit quality and on the ghost probability of the track associated to a hit can reduce the noise hit distribution significantly.

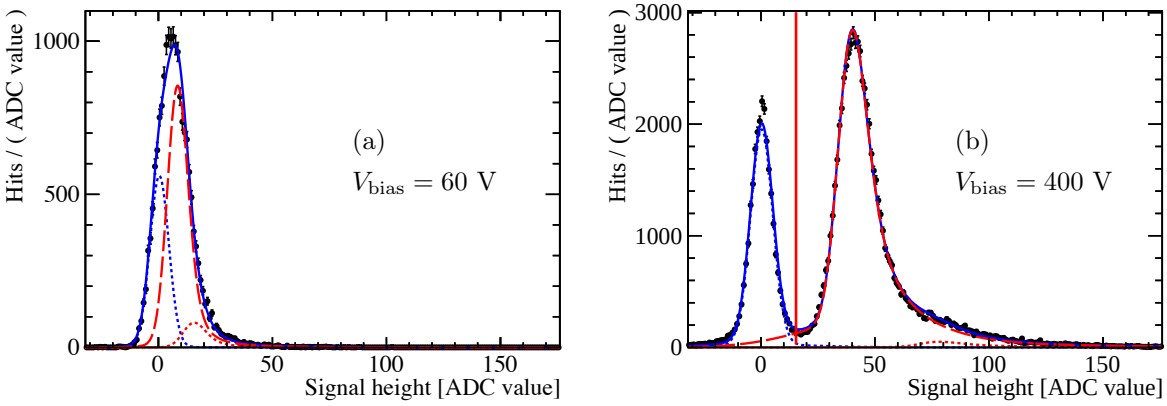


Figure 5.13: Signal height distribution of hits in TTaU associated to tracks before the quality requirements described in Fig. 5.14: (a) $V_{\text{bias}} = 60$ V, (b) $V_{\text{bias}} = 400$ V. The fit model is described in the text. The red vertical line shows the separation between tracks associated to noise and signal hits.

These track requirements are determined by using the amplitude distribution for $V_{\text{bias}} = 400$ (300) V in the TT (IT) and defining all tracks with an associated amplitude below 12 (20) ADC counts as either ghost tracks or tracks leading to bad extrapolations and those with an ADC count above 12 (20) as good tracks. Figure 5.14 shows the ratio between the number of signal and background tracks, S/B, as a function of the track χ^2/ndf and the track ghost probability. To reduce noise hits, only hits associated to tracks with values of track χ^2/ndf and track ghost probability below the white lines in Fig. 5.14 are used in the further analysis. The signal efficiency and background rejection efficiency of these requirements are shown in Fig. 5.15. These quantities are determined from the yields of signal and noise hits obtained from a fit of the amplitude distribution for $V_{\text{bias}} = 400$ (300) V in the TT (IT) with and without these requirements. The signal amplitude distributions including these requirements are displayed in Fig. 5.16 showing a significant background reduction.

5.3.3 Estimation of the Most Probable Value of the Signal Height

The most probable value (MPV) of the signal height per voltage and timing step is estimated from a maximum likelihood fit of the corresponding signal height distribution (*cf.* Fig. 5.18). To estimate the resolution in this distribution and the shape of the residual noise distribution after the track quality requirements, dedicated samples of noise hits for each read-out sector have been collected for each voltage and timing step. These noise hit samples are based on the summed ADC values of three (seven) random neighbouring strips in events where no reconstructed track is extrapolated to the corresponding TT (IT) read-out sector.
Figure 5.17 shows examples of these noise distributions for the TT and IT. As expected the

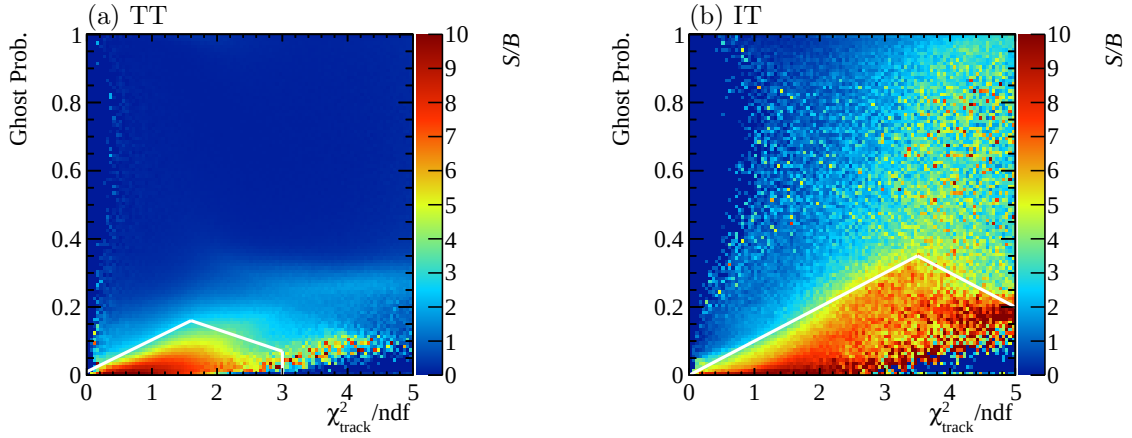


Figure 5.14: The observed ratio of the number of signal tracks to the number of background tracks as defined in the text, S/B, as a function of the track ghost probability and χ^2 (a) for the TT and (b) for the IT. The areas below the white lines correspond to the applied track requirements.

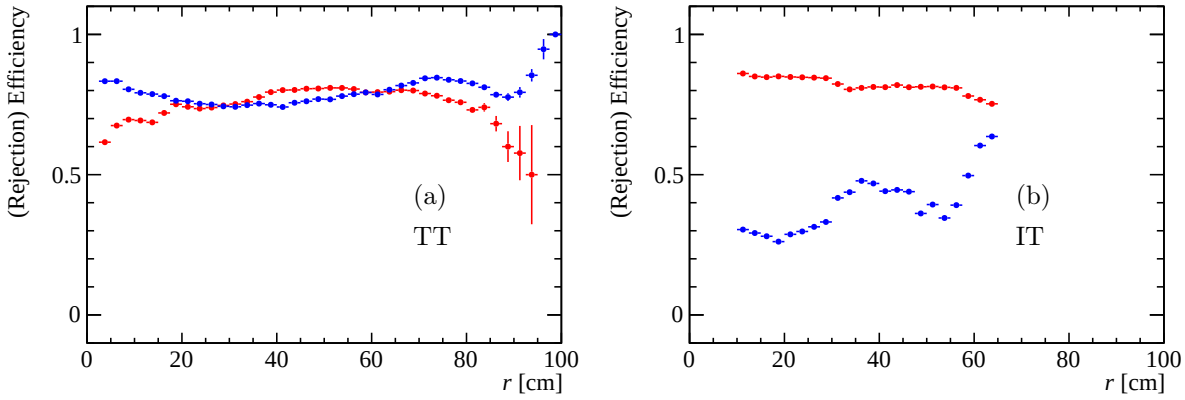


Figure 5.15: Efficiency for signal (blue) and rejection efficiency for background (red) of track samples in the CCE scans after the track requirement cut as a function of the distance, $r = \sqrt{x^2 + y^2}$, from the beam axis (a) for the TT and (b) for the IT.

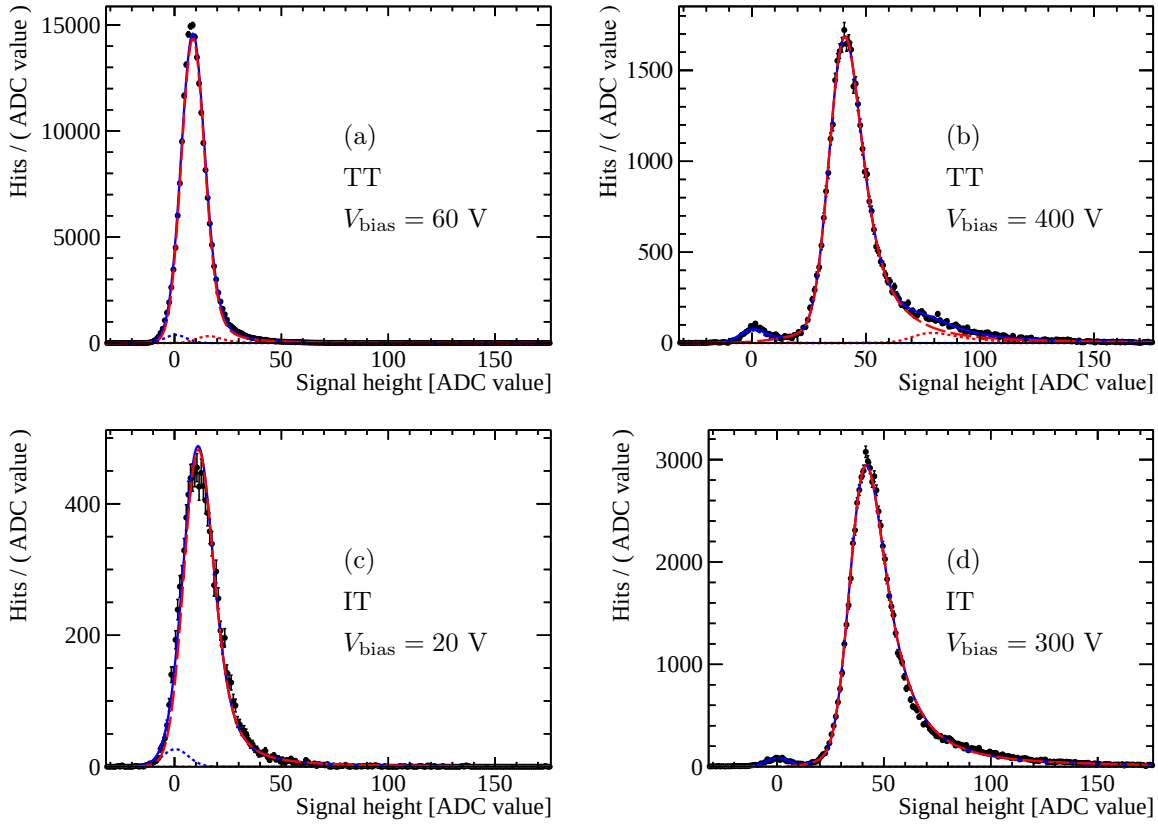


Figure 5.16: Signal height distribution for hits associated to tracks fulfilling the track χ^2 and ghost probability requirement described in Fig. 5.14. TTaU: (a) $V_{\text{bias}} = 60$ V, (b) $V_{\text{bias}} = 400$ V; T3X2: (c) $V_{\text{bias}} = 20$ V, (d) $V_{\text{bias}} = 300$ V. The fit model is described in the text.

distribution for the IT has a standard deviation that is larger than for the TT by a factor slightly smaller than $\sqrt{7/3}$, which would be the factor expected from the number of summed up strips. The reason for the slightly smaller factor is the smaller strip noise in IT. The measured distributions are described by a maximum likelihood fit of a double Gaussian

$$
\begin{aligned}
{}_d\mathcal{G}(x|\mu,\sigma_1,\sigma_2,f) &= f\cdot\mathcal{G}(x|\mu,\sigma_1)+(1-f)\cdot\mathcal{G}(x|\mu,\sigma_2) \\
&\text{with } \mathcal{G}(x|\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\exp\Big(-(x-\mu)^2/2\sigma^2\Big).
\end{aligned}
\tag{5.2}
$$

The central value μ is approximately zero. For TT, f is typically of the order of 0.98, σ_1 is about 4 ADC values and σ_2 about 16 ADC values, for the IT the values are 0.98, 6 ADC values and 23 ADC values, respectively.

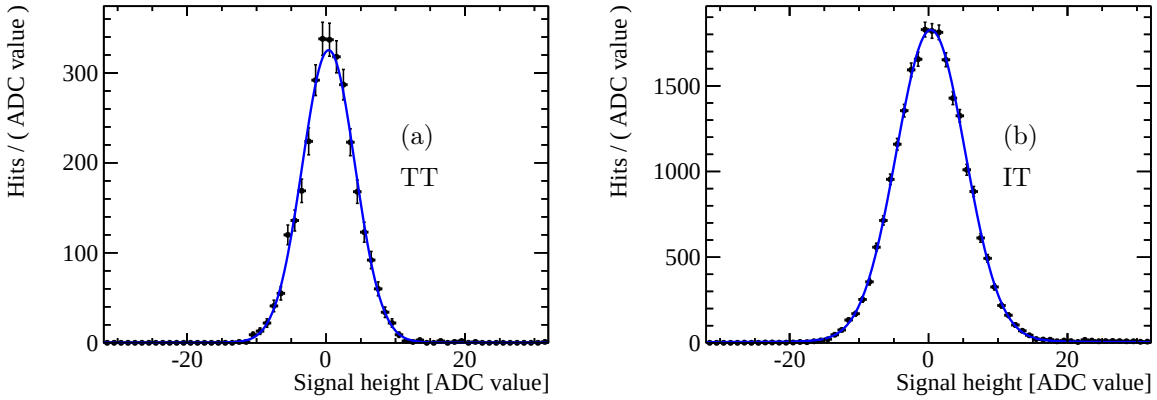


Figure 5.17: Signal height distribution of noise hits (a) in the TT and (b) in the IT. As the ADC values are summed over three strips in the TT and over seven strips in IT, the distribution for the IT is wider. The applied fit model is a double Gaussian.

The signal height distributions for hits associated to reconstructed tracks are described for the signal component by a convolution of a Landau distribution

$$
\mathcal{L}(x|m,s) = \lim_{\varepsilon\to 0}\frac{1}{2\pi i}\int_{\varepsilon-i\infty}^{\varepsilon+i\infty}\exp[s\ln(s)+(x+m)s], \tag{5.3}
$$

where m is the most probable value and s the spread, and the noise distribution described in Eq. (5.2) with fixed parameters f, μ, σ_1 and σ_2 from the fit to the dedicated noise sample. The contribution from noise hits is also described by Eq. (5.2) with fixed parameters from the fit to the noise hit sample.
For TT, a second Landau distribution with an MPV $2m$ and spread $2s$ is added. The reason are photons that convert into electron-positron pairs (*photon conversion*) in the material of the VELO. These electrons and positrons can be still reconstructed as long tracks if they create at least hits in three VELO stations, and have a very small opening angle. Therefore they generate signals in the same read-out strips in the TT and hence a signal height is detected that is twice as high as that from a single track.
Due to their opposite charge, their trajectories are bent into opposite directions by the dipole

magnet and therefore they do not generate signals in the same read-out strips in the IT (*cf.* Fig. 5.10(b)).
The smaller fraction of noise hits present in the IT sample is due to the larger number of summed up strips. Thus it is more likely that the strip collecting the largest amount of the charge carriers is included in the sum despite the less precise extrapolation as discussed above[4].
An example of all signal height distributions with $\delta t = 0$ ns in a specific CCE scan and for a particular read-out sector in the TT and IT can be found in Appendix A in Fig. A.1 and A.2, respectively.

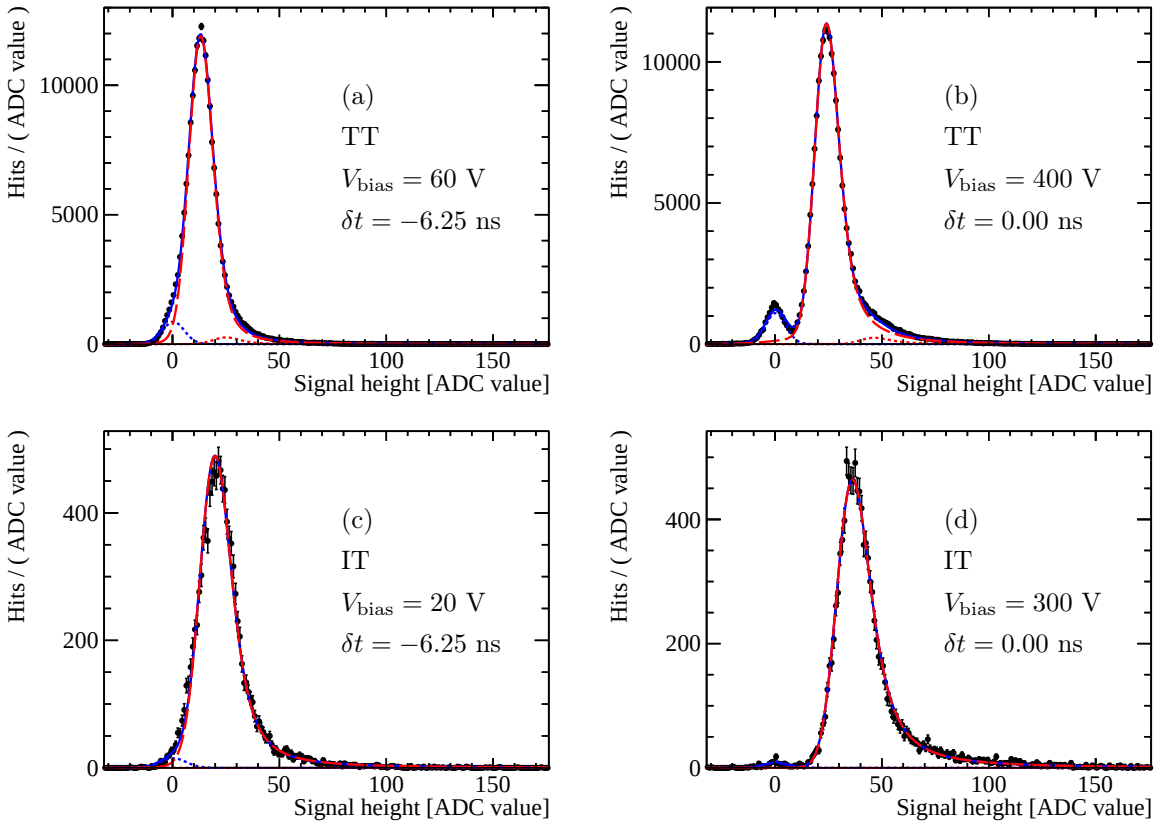


Figure 5.18: Signal height distribution for hits associated to tracks. TTaU: (a) $V_{\text{bias}} = 60$ V, $\delta t = -6.25$ ns, (b) $V_{\text{bias}} = 400$ V, $\delta t = 0.00$ ns; T3X2: (c) $V_{\text{bias}} = 20$ V, $\delta t = -6.25$ ns, (d) $V_{\text{bias}} = 300$ V, $\delta t = 0.00$ ns. The data is fit with a double Gaussian describing the residual noise hits (blue dotted line) and a convolution of the double Gaussian distribution with a Landau distribution for the signal hits (red dashed line). In the case of TT, a second Landau distribution convolved with the double Gaussian distribution (red dotted line) is added to describe the contribution from merged electron-positron pairs due to photon conversion in the VELO.

[4] Summing up more strips is in the TT not necessary as shown by the signal height distributions. But signal height distributions in the IT only based on the summed ADC values of three strips are heavily dominated by noise hits.

5.3.4 Estimation of the Amount of Collected Charge

As the MPVs in each Bettle delay time δt and voltage step are determined from the fit described above, the pulse shape as a function of δt in each voltage step can be estimated from a fit of the MPVs for the different values of δt. The pulse shape can be described by the empirical function

$$f(\delta t|\tau, \delta t_0, A) = A \cdot \exp\left(-\frac{\delta t - \delta t_0}{\tau}\right) \cdot \left[\frac{1}{2}\left(\frac{\delta t - \delta t_0}{\tau}\right)^2 - \frac{1}{6}\left(\frac{\delta t - \delta t_0}{\tau}\right)^3\right], \qquad (5.4)$$

which is motivated by the use of a CR-RC shaper in the Beetle chip[5]. This function describes the pulse shape well between its two zeros ($\delta t = \delta t_0$ and $\delta t = \delta t_0 + 3 \cdot \tau$) [136].
Figure 5.19 shows examples of this fit for different bias voltage settings in the TT and IT. From the fitted function, the integral between δt_0 and $\delta t_0 + 3 \cdot \tau$ is determined. This integral divided by 20 ns is referred as Charge Equivalent in the following to estimate the amount of collected charge[6]. Figures A.3 and A.4 in the Appendix A show as an example all the obtained pulse shapes for the different bias voltage settings in a CCE scan for selected read-out sectors in the TT and IT .

5.3.5 Extraction of the Depletion Voltage

Examples of the Charge Equivalent as a function of V_{bias} are shown in Fig. 5.21 for a given read-out sector and CCE scan. As there are large uncertainties on the data points for the lowest values of V_{bias} (TT: $V_{\text{bias}} = 60$ V, IT: $V_{\text{bias}} = 20$ V) due to the overlap of the signal and noise amplitude distributions, these data points are discarded. For the IT also the data point at $V_{\text{bias}} = 300$ V is removed as there are several instances where the corresponding data were collected with incorrect bias voltage or delay time settings.
In contrast to the capacitance in a C-V measurement, the Charge Equivalent approaches its maximal value asymptotically as the electric field in the silicon increases with increasing bias voltage, leading to a better charge collection efficiency even above V_{depl}. Thus, the depletion voltage is not as well defined by the characteristic of the Charge Equivalent as a function of V_{bias} as it would be from a C-V scan.
The depletion voltage, V_{depl}, is determined from a fit to the data points. As baseline for the fit a third-order spline, $S(V_{\text{bias}})$, is used. It has been chosen as its form in the region of $V_{\text{bias}} \approx V_{\text{depl}}$ is insensitive to statistical fluctuations for data points at V_{bias} much smaller or larger than V_{depl}. For the spline it is required that $S(0) = 0$, and $\mathrm{d}S/\mathrm{d}V_{\text{bias}} = 0$ for $V_{\text{bias}} = 500$ (250) V in the TT (IT). The value of $S_{\text{max}} = S(V_{\text{bias}})$ for $V_{\text{bias}} = 500$ (250) V in the TT (IT) is determined from a least square fit of a constant to the data points $V_{\text{bias}} > 250$ (125) V.
The earliest CCE scan listed in Tab. 5.3 is used to calibrate the extraction of V_{depl} from the splines. The ratio between the value of $S(V_{\text{depl}})$, where V_{depl} is the depletion voltage measured in C-V scans after the production of the sensors[7] and S_{max} is calculated (*cf.* Fig. 5.20(a) and (c)) for each read-out sector. Figures 5.20(b) and (d) show the distribution of this ratio for read-out sectors in the TT and IT, respectively. In the IT, all read-out sectors are used in this

[5] In the read-out electronics, two CR-elements are followed by two RC-elements.
[6] The division by 20 ns ($\approx 1.5\tau$) is applied that the Charge Equivalent has a similar value as the pulse height.
[7] As the read-out sectors mostly consist of several sensors, the mean of the measured depletion voltage of the different sensors is taken. The difference in V_{depl} among the sensors is at maximum 10 V as sensors with similar depletion voltage values have been grouped in the same read-out sector.

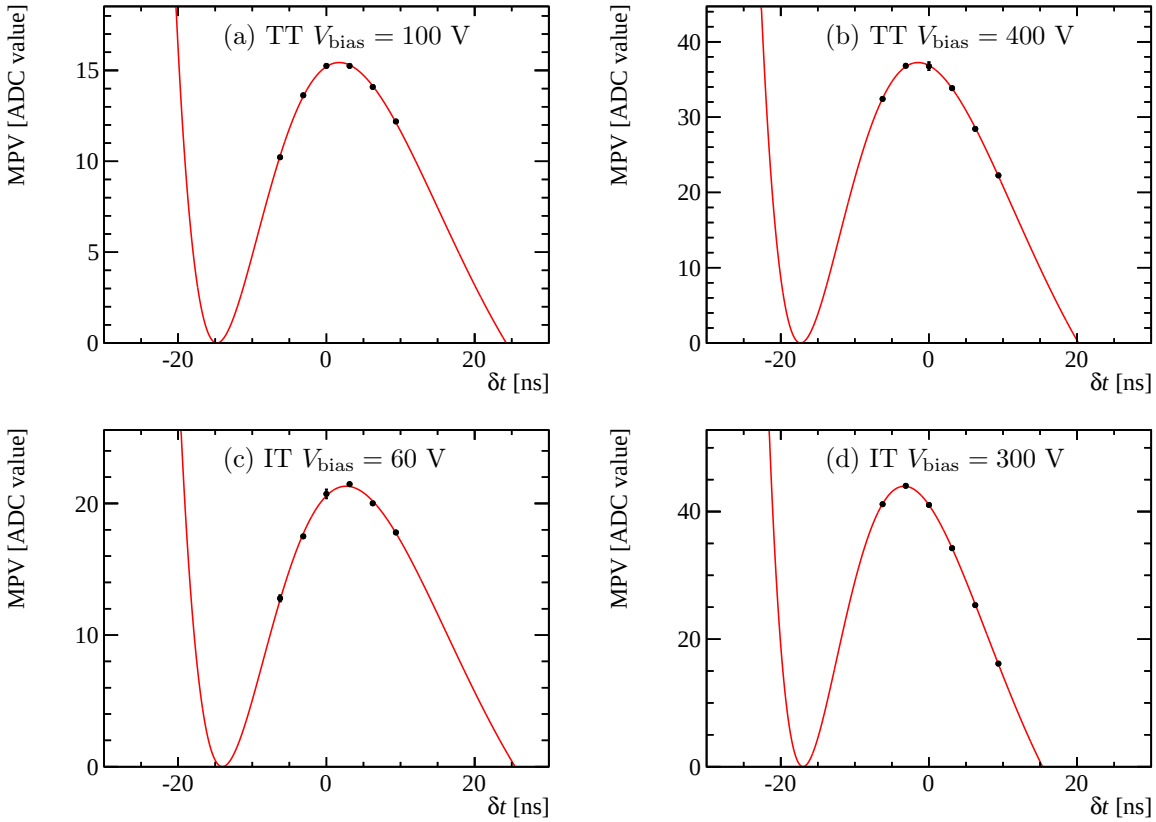


Figure 5.19: Pulse shapes of timing scans in the TT and IT for different bias voltage values, fit with the function described in Eq. (5.4). TT: (a) $V_{\text{bias}} = 100$ V, (b) $V_{\text{bias}} = 400$ V; IT: (c) $V_{\text{bias}} = 60$ V, (d) $V_{\text{bias}} = 300$ V.

determination while in the TT only the read-out sectors in the TTaU layer that are part of a two-sector half-module (*i.e.* the half-module consisting of a three- and a four-sensor read-out sector, *cf.* Fig. 3.6) are taken. In the region of those half-modules only a very small irradiation up to the time of the first CCE scan took place. The mean values, m_r of the obtained ratio distributions for the TT and IT are taken as fraction of the maximal Charge Equivalent $S_{\max}$ corresponding to the Charge Equivalent of the depletion voltage, *i.e.* V_{depl} s.t. $S(V_{\text{depl}}) = m_r S_{\max}$. Their standard deviation of the distributions is taken as uncertainty on these values leading to $m_r = (95.3 \pm 1.9)\,\%$ for the TT and $m_r = (95.6 \pm 1.7)\,\%$ for the IT.
Figure 5.21 illustrates the extraction of V_{depl} using these percentages. Figures 5.21(a) and (b) as well as (c) and (d) also show a decrease in the total charge collection efficiency as the maximal Charge Equivalent value decreases over time. This is – as discussed in the previous chapter – an effect of the irradiation. The figures show even that the charge collection efficiency is in the read-out sector part closer than 45 mm to the beam axis (*i.e.* the region with the highest radiation) as expected smaller than over the whole read-out sector. The opposite behaviour is visible for the IT (*cf.* Figs 5.21(e) and (f)), which is not yet fully understood.

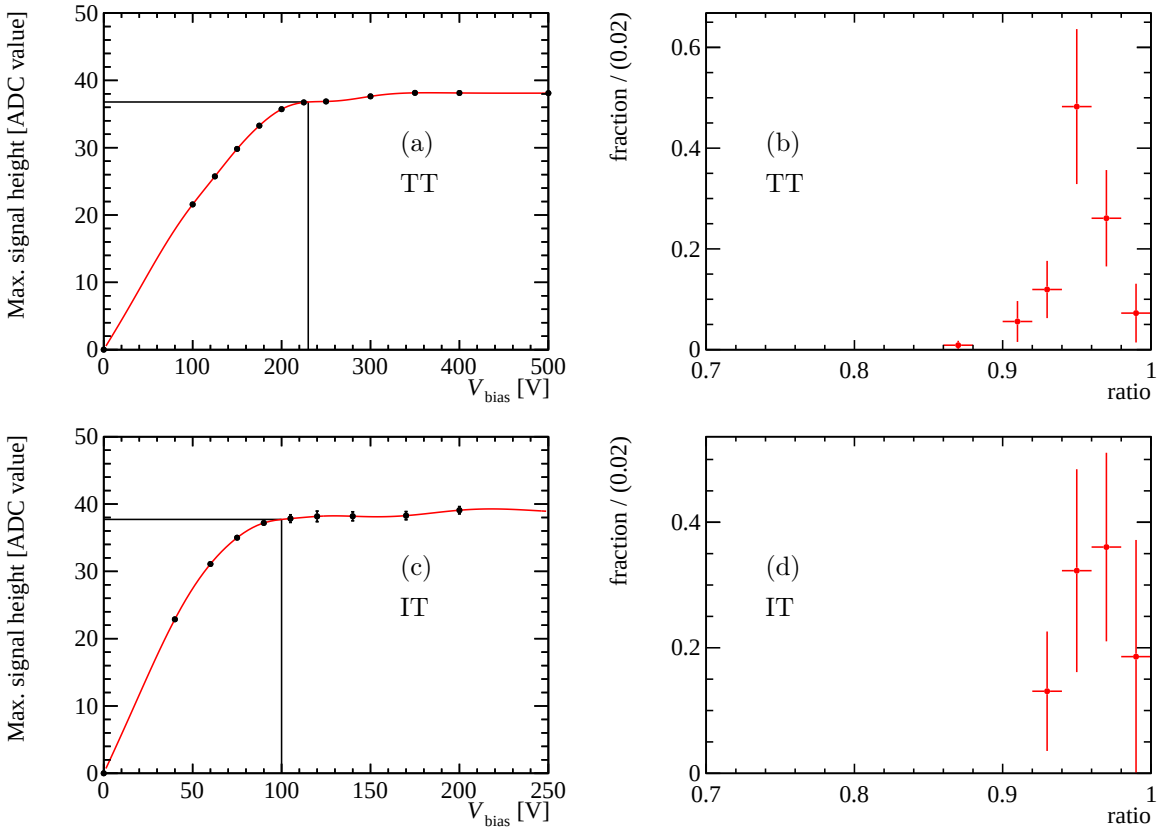


Figure 5.20: Extraction of the ratio, m_r between the Charge Equivalent corresponding to the measured value of V_{depl} after production and the maximal Charge Equivalent (a) in the TT and (c) in the IT. The distribution of m_r over all the considered sensors is shown in (b) for the TT and in (d) for the IT.

The obtained values for V_{depl} have also been checked for systematic uncertainties:

1. **Track momentum** The values of V_{depl} have been determined separately for hits associated

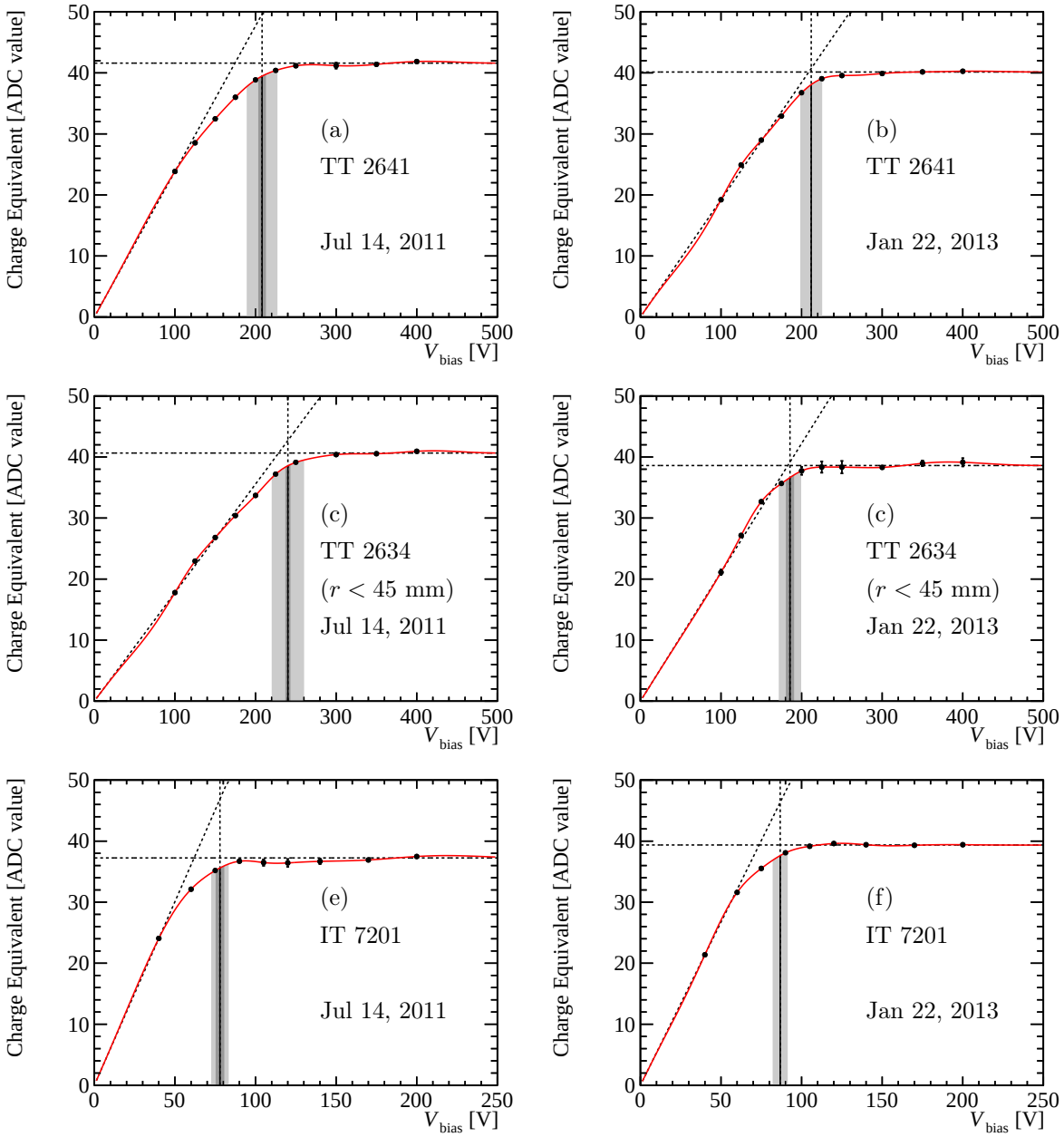


Figure 5.21: Charge Equivalent (pulse shape integral) as a function of the applied bias voltage for the CCE scans from July 7, 2011 (Fill 1944, left column) and January 22, 2013 (Fill 3478, right column). The data points are fit with a third-order spline. The depletion voltage V_{depl} is extracted as the bias voltage at which the fitted spline reaches a certain fraction (about 95 %, *cf.* text) of its maximum value. The vertical lines show the obtained values of V_{depl}. The dark grey bands show the statistical uncertainties, the lighter grey bands the systematic uncertainties. The dashed horizontal lines show the fitted maximal Charge Equivalents and the diagonal ones the derivatives of the spline at $V_{\mathrm{bias}} = 0$ V.
The displayed plots correspond to (a,b) the TT read-out sector 2641 (two sensors), (c,d) the inner part ($r < 45$ mm) of TT read-out sector 2634 (one sensor) and (e,f) the IT read-out sector 7201 (two sensors). While sectors 2641 and 7201 show no significant decrease in V_{depl}, there is a significant decrease in V_{depl} between the two CCE scans in sector 2634.

to tracks with $p > 30\,\mathrm{GeV}/c$ and $p < 30\,\mathrm{GeV}/c$. The on average obtained difference between the low- and high-momentum sample is smaller than the statistical uncertainties on the results.

2. **Track quality** The values of V_{depl} were determined by applying cuts of $\chi^2_{\mathrm{track}}/\mathrm{ndf}$ smaller than 1 and of ghost probability smaller than 0.1 instead of the cuts shown in Fig. 5.15. The obtained estimations have larger statistical uncertainties due to the smaller samples, but are within the uncertainties in agreement with the nominal results.

3. **Fit of the signal height distribution** The signal height distributions were fit for the IT as well as for the TT with the model consisting of a double Gaussian (background) and a convolution of this double Gaussian with a Landau distribution (signal). This allows to check the impact of the description of photon conversion for TT data. The fit range was constraint with an upper boundary 1.5 times larger than the modal value of the data distribution. The obtained values of V_{depl} using this fit model are within statistical uncertainties also in agreement with the nominal results.

4. **Pulse shape** Instead of using the Charge Equivalent as integral of the fitted pulse shape, the maximal value of the pulse shape as a function of V_{bias} was used to extract V_{depl}. Also in this case there was no significant systematic difference spotted.

5. **Functional form of the Charge Equivalent as a function of V_{depl}** This effect is estimated by using an alternative method to extract V_{depl} and taking the difference to the nominal value as assigned systematic uncertainty. In the alternative method, the data points are fitted by the phenomenological function

$$S(V_{\mathrm{bias}}) = \begin{cases} \dfrac{A_0}{1+\exp\left(r\frac{V_{\mathrm{bias}}-V_0}{V_{\mathrm{depl}}-V_0}\right)} & \text{for } V_{\mathrm{bias}} \geq V_0 \\ \frac{A_0}{2}[1+\frac{r}{2}(V_{\mathrm{bias}}-V_0)] & \text{for } V_{\mathrm{bias}} < V_0, \end{cases} \tag{5.5}$$

   where r is chosen such that $S(V_{\mathrm{depl}}) = m_r S_{\mathrm{max}}$. The values of m_r for the TT and IT are estimated as described above, but using the function in Eq. (5.5) instead of the third-order spline.

6. **Variation of m_r** The last systematic uncertainty is due to the standard deviation, σ_{m_r}, of $S(V_{\mathrm{depl}})/S_{\mathrm{max}}$ based on the third-order spline. The average absolute difference between V_{depl} extracted based on the spline and the bias voltage values such that $V_{\mathrm{bias}} = S^{-1}((m_r \pm \sigma_{m_r}) \cdot S_{\mathrm{max}})$ is taken as systematic uncertainty.

The last two systematic uncertainties are the only relevant ones while the first four systematic uncertainties are negligible as described above. Thus, the last two uncertainties are added in quadrature to give the total systematic uncertainty for each measured depletion voltage.

5.3.6 Results and Comparison between Data and Simulation

Figure 5.22 shows the determined values of V_{depl} for the different read-out sectors and CCE scans as a function of the average 1-MeV neutron equivalent fluence in the corresponding sector integrated up to the time of the corresponding CCE scan. The fluences are estimated from the delivered luminosity as a function of time and the FLUKA simulation.

In the case of the TT, V_{depl} is also determined separately for the innermost regions of the six one-sensor read-out sectors closest to the beam pipe. In these measurements, only tracks that cross the sensors within a certain radial distance, r, from the beam axis are considered. The chosen distances are $r = 45$ and 75 mm.

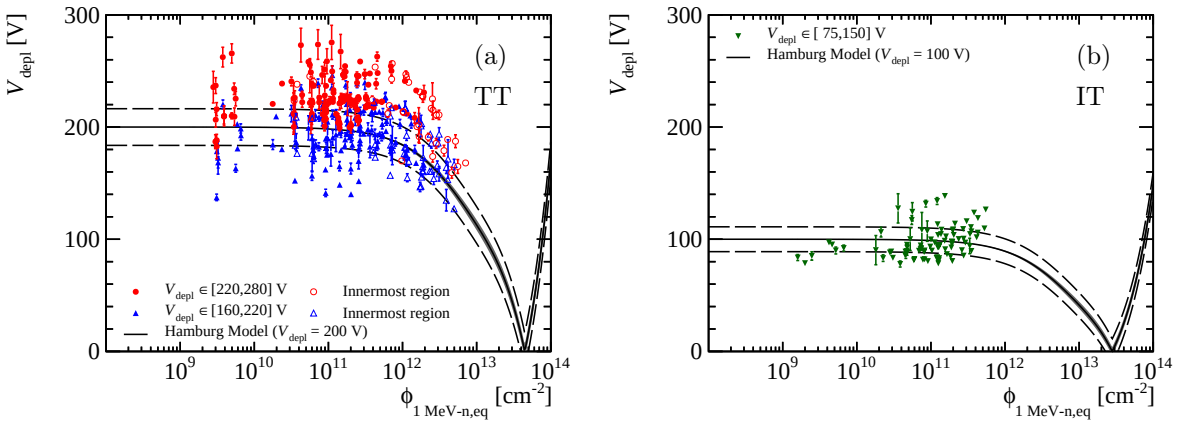


Figure 5.22: Measured values of V_{depl} from the different CCE scans and the different read-out sectors (a) in the TT and (b) in the IT as a function of the 1-MeV neutron equivalent fluence obtained from the running conditions and FLUKA. The error bars of the data points display the statistical uncertainty. The solid black curve shows the predictions based on the stable damage part of the Hamburg model, the grey shaded region its uncertainty due to the parameter uncertainty of the model. The dashed black lines show the average systematic uncertainty on the measured V_{depl} values. The hollow data points show the measurements in the TT considering only reconstructed tracks that traversed the read-out sectors closest to the beam pipe in a distance closer than 45 or 75 mm.

Also shown in Fig. 5.22 are predictions of the evolution of V_{depl} as a function of $\Phi_{1\text{ MeV-n,eq}}$, based on the stable-damage part of the Hamburg model (*cf.* Eq. (4.17)). The initial V_{depl} value in the prediction is 200 V in the TT and 100 V in the IT. As predicted by the simulation, a decrease of the measured V_{depl} values as a function of the fluence is visible in the TT. No such decrease has been measured in the IT. Several IT read-out sectors show a slightly increasing behaviour. This might be in connection with the observed increase of S_{max} as described before. Nevertheless, the measured values are within their uncertainties in agreement with the predictions. As the effects of radiation damage is more important in the TT, the studies have been also concentrated on this sub-detector, where a very good agreement between the observed change in V_{depl} and the Hamburg model predictions is observed.

For selected read-out sectors, the expected evolution of V_{depl} as a function of time is calculated using the full Hamburg model as described in Sec. 4.3.2. This calculation is based on the predicted fluence from the FLUKA simulation, the actual running conditions including the measured temperatures in the detector boxes and the initial values of V_{depl} measured after production. The initial values of the depletion voltage have been measured to a precision of 5 V in the C-V scans. Examples of these predictions together with the V_{depl} values are shown in Figs 5.23 and 5.24. Predictions and measurements from the CCE scans agree well within uncertainties. Further examples can be found in Appendix A.

The used parameters of the Hamburg model are listed in Tab. 5.4. The uncertainty on the predictions is defined as the 68 % confidence interval with central coverage due to the uncertainties on the model parameters, on the FLUKA simulation and on the temperature measurements. All

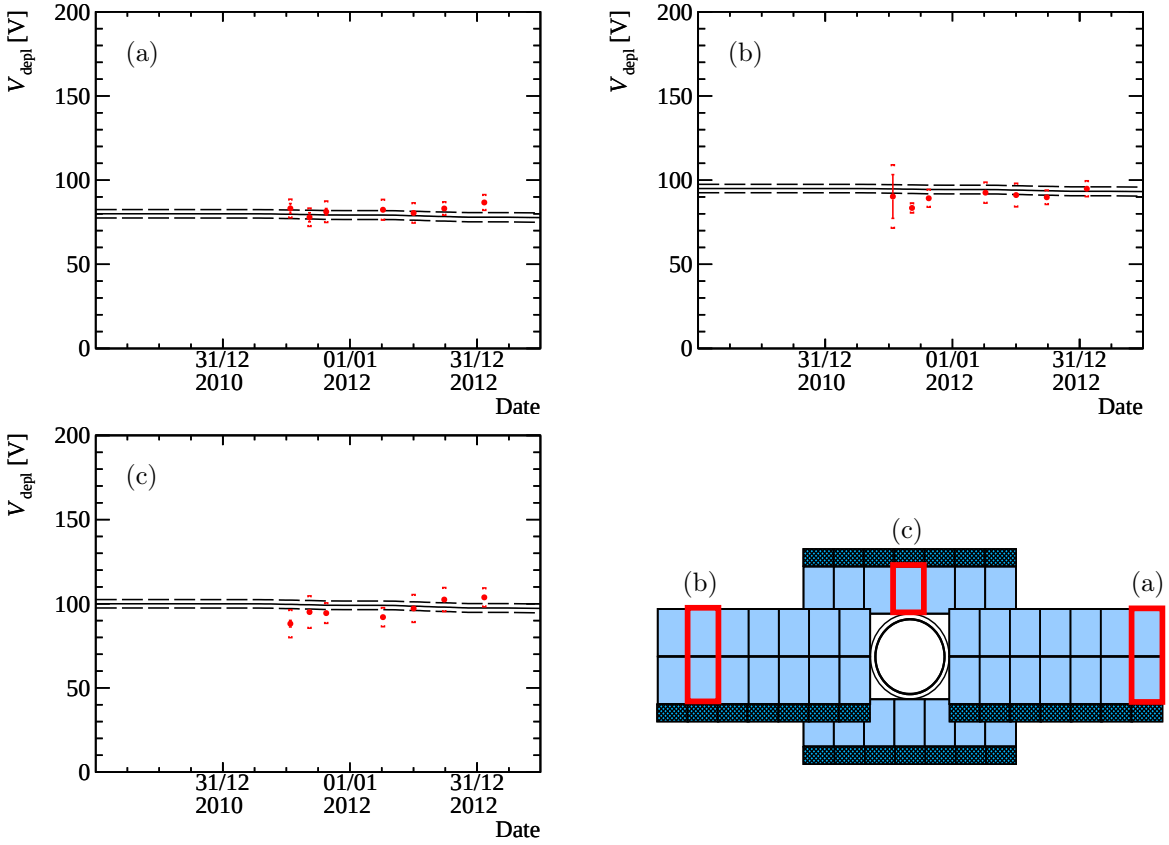


Figure 5.23: Evolution of V_{depl} as a function of time for the IT read-out sectors (a) 7201, (b) 7238, and (c) 7301. The red points show the measurements based on the CCE scans. The scan from the early fill 1616 is excluded as it was used for the calibration of V_{depl} as described in the text. The inner error bars show the statistical uncertainties, the outer ones the statistical and systematic uncertainties added in quadrature. The black line shows the prediction from the Hamburg model based on the actual data taking conditions, temperature measurements and FLUKA simulation. The grey area shows the uncertainty on the prediction due to the uncertainties of the model parameters, of the temperature measurement and of the FLUKA simulation. The dashed lines show the ± 5 V uncertainty of the initial V_{depl} measurement after sensor production. As expected there is no significant decrease in V_{depl} up to the end of Run I in the IT.

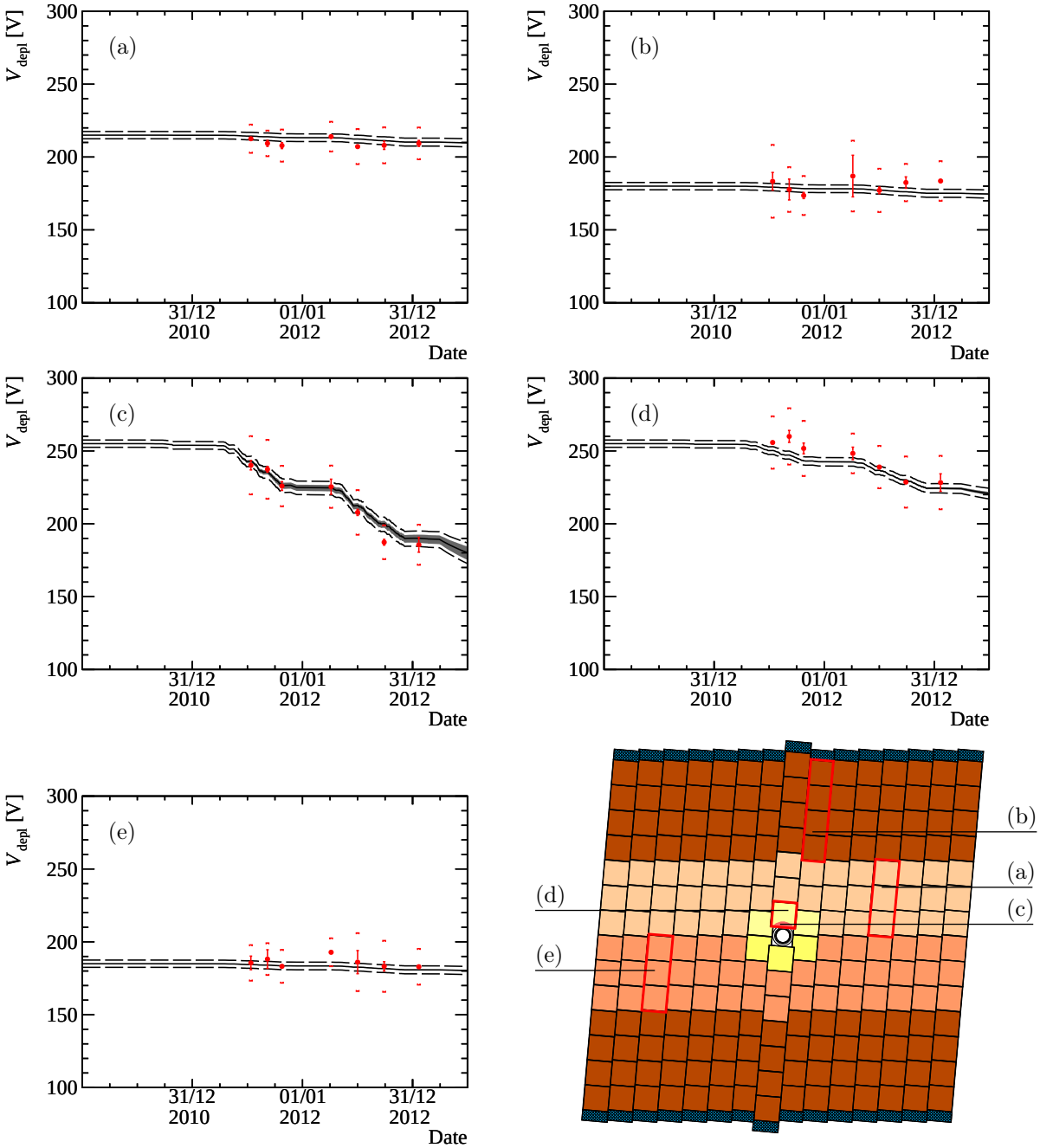


Figure 5.24: Evolution of V_{depl} as a function of time for TT read-out sectors (a) 2607, (b) 2630, (c) 2634, (d) 2634 ($r < 45$ mm) and (e) 2670. The red points show the measurements based on the CCE scans. The scan from the early fill 1616 is excluded as it was used for the calibration of V_{depl} extraction as described in the text. The inner error bars show the statistical uncertainties, the outer ones the statistical and systematic uncertainties added in quadrature. The black line shows the prediction from the Hamburg model based on the actual data taking conditions, on the temperature measurements and on the FLUKA simulation. The grey band shows the uncertainty on the prediction due to the uncertainties of the model parameters, of the temperature measurement and of the FLUKA simulation. The dashed lines show the ± 5 V uncertainty of the initial V_{depl} measurement after sensor production. While V_{depl} does not decrease significantly in the outer region of the detector layer, there is a significant decline of V_{depl} in the innermost region.

these uncertainties are treated as Gaussian uncertainties.
Figure 5.25 shows the change in V_{depl} up to the second (July 2011) and the last CCE scan (January 2013) for the innermost region of TTaU. A significant change in V_{depl} is visible for the innermost region.

Table 5.4: Parameters for the depletion voltage simulation by the Hamburg model.

Parameter	Value
$n_{c,0}$	$(3.28 \pm 0.26) \times 10^{-10} \text{cm}^{-3}$
c	$2.29) \times 10^{-13} \text{cm}^{2}$
g_c	$(1.60 \pm 0.04) \times 10^{-2}$ cm^{-1}
g_a	$(1.40 \pm 0.14) \times 10^{-2}$ cm^{-1}
g_r	$(5.70 \pm 0.09) \times 10^{-2}$ cm^{-1}
$k_{a,0}$	$(2.4 \pm 1.0) \times 10^{15}$ s^{-1}
$k_{r,0}$	$(1.5 \pm 1.1) \times 10^{15}$ s^{-1}
E_{aa}	(1.09 ± 0.03) eV
E_{ar}	(1.31 ± 0.03) eV

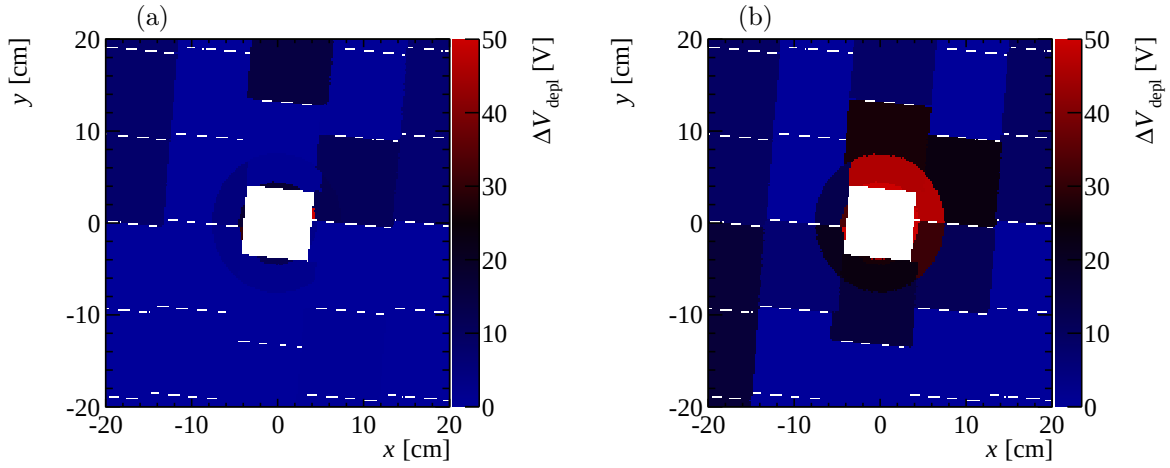


Figure 5.25: Absolute change in V_{depl} in the innermost region of TTaU up to the CCE scan (a) on July 7, 2011 and (b) on January 22, 2013. A significant decrease between the two CCE scans is visible in the sensors closest to the beam line.

5.4 Interpretation of the Results

From the measurements of the leakage current and the depletion voltage in the ST, it can be concluded that the radiation damage of the detector follows closely expectations. A type inversion of the sensors (*i.e.* the sensor bulk starts to become *p*-type-like), even in the innermost part of TT, is not expected up to the collection of about ten times more luminosity including the fact

that the fluence is expected to increase with the increasing centre-of-mass energy from 8 TeV to 13-14 TeV (*cf.* Fig. 5.22(a)) in LHC Run II. The planned amount of data to be collected by LHCb in Run II (2015-2018) corresponds to 5 fb^{-1} at $\sqrt{s}$ = 13-14 TeV. The sub-detectors of the ST should thus withstand without problems the irradiation until the second long shutdown (LS2), in which the detectors are foreseen to be replaced (the TT by the Upstream Tracker (UT), designed also as silicon micro-strip detector; the IT and the OT by a tracker made of scintillating fibres [8]) as part of the upgrade campaign of LHCb.
Nevertheless, it is crucial to keep the detector boxes cold also in the periods of technical stops and winter shutdowns due the accelerated reverse annealing effects at higher temperatures and to conserve the good performance of the sensors.

III

The Search for the Rare Decays $B^0_{(s)} \to \mu^+\mu^-$

The path from error to truth
is rare and beautiful!

La légende des siècles, préface, Victor Hugo (1802 - 1885)

6

The Decays $B^0_{(s)} \to \mu^+\mu^-$ and Their Analysis in LHCb

The search for the decays $B^0_{(s)} \to \mu^+\mu^-$ and the measurement of their branching fractions is one of the key analyses in LHCb as these rare decays serve as indirect probes for physics beyond the SM (BSM). Their theoretical implications are discussed in Sec. 6.1 while the experimental status before the start of the LHC is discussed in Sec. 6.2. The analysis strategy for this search in LHCb is described in Sec. 6.3.

6.1 Theoretical Implications of $B^0_{(s)} \to \mu^+\mu^-$

The decays $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ involve transitions of type $b \to s$ and $b \to d$, respectively. Such processes, where a quark changes its flavour, but not its charge, are called Flavour Changing Neutral Currents (FCNC). FCNC are forbidden in the SM at tree level by the GIM-mechanism [155] and are only allowed either via Penguin or via Box diagrams (*cf.* Fig 6.1) at lowest order. Therefore they are highly suppressed in the SM . Their large suppression allows that contributions from BSM might be of the same order as the SM contribution, leading to significant changes in observables and therefore making decays involving FCNC promissing places to search for effects from BSM. In addition to the suppression as FCNC the decays $B^0_{(s)} \to \mu^+\mu^-$ are also suppressed in the SM by a factor $(2m_\mu/m_{B^0_{(s)}})^2$ due to helicity.

6.1.1 $B^0_{(s)} \to \mu^+\mu^-$ in the Standard Model

The branching fractions $B^0_{(s)} \to \mu^+\mu^-$ can be expressed in a HQEFT by the Wilson coefficients C_{10} (axial-vector coupling), C_S (scalar coupling) and C_P (pseudo-scalar coupling) and their

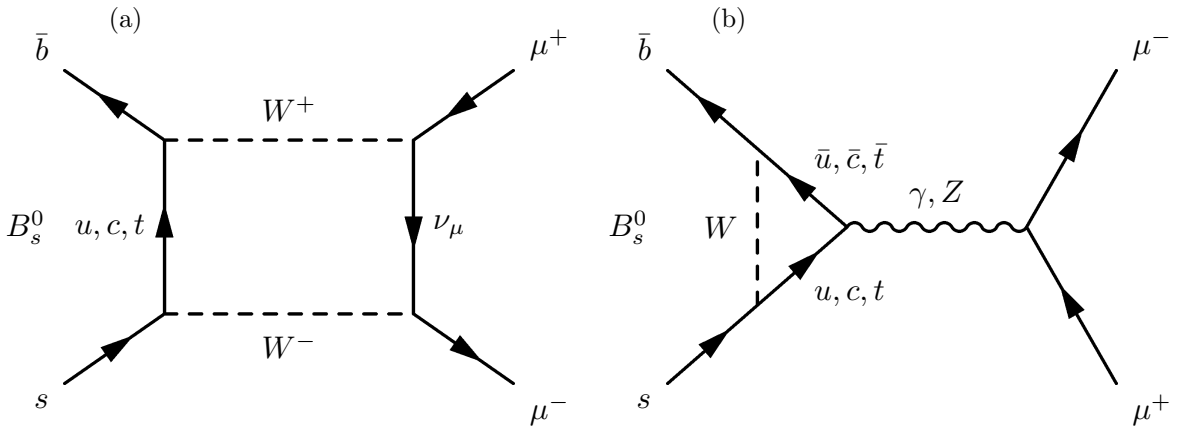


Figure 6.1: Feynman diagrams for the decay $B^0_s \to \mu^+\mu^-$ in the SM: (a) Box diagram; (b) Penguin diagram

opposite-helicity counterparts [83] as

$$\begin{aligned}\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-) &= \frac{G_F^2 \alpha^2 m^3_{B^0_{(s)}} f^2_{B^0_{(s)}} \tau_{B^0_{(s)}}}{64\pi^3 \sin^4\theta_W} \cdot |V_{tb}V^*_{ts/d}|^2 \left(1 - \frac{4m_\mu^2}{m^2_{B^0_{(s)}}}\right)^{1/2} \cdot \\ &\quad \left[\left(1 - \frac{4m_\mu^2}{m^2_{B^0_{(s)}}}\right)|C_S - C'_S|^2 + \right. \\ &\quad \left. \left|C_P - C'_P + \frac{2m_\mu}{m_{B^0_{(s)}}}(C_{10} - C'_{10})\right|^2\right]. \end{aligned} \tag{6.1}$$

The parameter G_F is the Fermi constant, α the fine-structure constant, θ_W the Weinberg angle, m_μ the mass of the muon, $m_{B^0_{(s)}}$ the mass of the B^0 or B^0_s meson, $\tau_{B^0_{(s)}}$ its lifetime, and $f_{B^0_{(s)}}$ its decay constant.

The only contribution allowed in the SM comes from the axial-vector coupling C_{10} while all other couplings are either forbidden ($C^{(\prime)}_{S,P}$) [1] or heavily suppressed by a factor $m_\mu/m_{B^0_{(s)}}$ (C'_{10}). The SM prediction for C_{10} can be calculated at leading order [92, 117] by

$$C_{10} = \frac{x}{8}\left(\frac{4-x}{1-x} + \frac{3x}{(1-x)^2}\ln x\right), \text{ where } x = \frac{m_t^2(\mu)}{m_W^2} \tag{6.2}$$

and m_W is the mass of the W boson and $m_t(\mu)$ is the top-quark mass renormalised in the $\overline{\text{MS}}$ scheme.

The CP-averaged SM predictions for the branching fractions [97] are

$$\begin{aligned}\mathcal{B}(B^0_s \to \mu^+\mu^-) &= (3.35 \pm 0.28) \times 10^{-9} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &= (1.07 \pm 0.10) \times 10^{-10},\end{aligned}$$

[1] There is a scalar contribution present in the SM, which is mediated by the Higgs boson, but is negligible due to the small Yukawa couplings of the muons. Other scalar or pseudo-scalar contributions do not exist in the SM due to the lack of corresponding intermediate particles.

which use the recent world average of $\tau_{B^0_s} = 1.516 \pm 0.011$ ps [59]. The largest contributions to the theoretical uncertainty come from $|V_{tb}V^*_{ts/d}|$ and the decay constant, $f_{B^0_{(s)}}$, which is estimated using lattice QCD [96][2].
The values of $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ quoted above are calculated for a decay time of the $B^0_{(s)}$ meson of $t = 0$. So far all experiments have measured the $B^0_{(s)} \to \mu^+\mu^-$ branching fractions in a time-integrated manner. Therefore it is required to correct the SM predictions for $B^0_s \to \mu^+\mu^-$ due to the non-negligible decay width difference, $\Delta\Gamma_s$, between the heavy and the light CP-eigenstates in the B^0_s-system [125].
The CP-averaged[3] ($\mathcal{B}(B^0_s \to \mu^+\mu^-)$) and the time-averaged ($\mathcal{B}(B^0_s \to \mu^+\mu^-)^{\langle t \rangle}$) branching fraction predictions are connected by

$$\mathcal{B}(B^0_s \to \mu^+\mu^-)^{\langle t \rangle} = \frac{1 + \mathcal{A}_{\mu\mu} \cdot y_s}{1 - y_s^2} \cdot \mathcal{B}(B^0_s \to \mu^+\mu^-), \tag{6.3}$$

where $\mathcal{A}_{\mu\mu} = (\Gamma_{B^0_{s,H} \to \mu^+\mu^-} - \Gamma_{B^0_{s,L} \to \mu^+\mu^-})/(\Gamma_{B^0_{s,H} \to \mu^+\mu^-} + \Gamma_{B^0_{s,L} \to \mu^+\mu^-})$ is the channel specific decay width asymmetry between the heavy and the light CP-eigenstates and $y_s = \Delta\Gamma_s/2\Gamma_s$. The values of these model-dependent parameters in the SM are $y_s = 0.0615 \pm 0.0085$ [59] and $\mathcal{A}_{\mu\mu} = 1$ since $B^0_{(s)} \to \mu^+\mu^-$ as two-body decays of pseudo-scalar mesons into fermions are pure CP-odd decays.
This leads to a SM prediction for the time-integrated branching fraction of

$$\mathcal{B}(B^0_s \to \mu^+\mu^-)^{\langle t \rangle} = (3.56 \pm 0.30) \times 10^{-9}.$$

Due to the smaller width difference in the B^0-system, the branching fraction of $B^0 \to \mu^+\mu^-$ is not significantly affected by this mechanism.
Recent studies including three-loop QCD corrections in the calculation of the SM value for C_{10} [168] and two-loop EW corrections in the calculations of the contributing operators [85] led to the most accurate prediction [84]

$$\mathcal{B}(B^0_s \to \mu^+\mu^-)^{\langle t \rangle} = (3.65 \pm 0.23) \times 10^{-9}.$$

6.1.2 $B^0_{(s)} \to \mu^+\mu^-$ beyond the Standard Model

Model-Independent Consideration

Many models in BSM predict a significant deviation of these branching fractions from their SM values. Some of these models are discussed in the paragraphs below.
The impact of physics beyond the SM can be described in a generic approach based on the Wilson coefficients. The presence of scalar or pseudo-scalar couplings can lead to a change in the branching fractions. Furthermore, modifications in the axial-vector coupling, described by C_{10}, as well as its helicity structure ($C'_{10} \neq 0$) can also change $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$. Equation (6.1) shows that there is always an enhancement of the branching fractions in presence of scalar

[2] As an alternative, the branching fraction can also be expressed as a function of $\Delta m_{B^0_{(s)}}$, the mass difference between the heavy and light weak eigenstates of the B^0-$\overline{B}^0$ or B^0_s-$\overline{B}^0_s$ meson system [96]. This allows to remove the CKM term from the calculation as well as trading the decay constant $f_{B^0_{(s)}}$ for a bag parameter $\hat{B}_{B^0_{(s)}}$, which has a smaller uncertainty [184].

[3] It corresponds to the branching fraction at $t = 0$ as an equal abundance of B^0_s and $\overline{B}^0_s$ mesons is assumed.

couplings while there can be a destructive interference between the pseudo-scalar and axial-vector contributions, leading to smaller branching fractions than in the SM. This might for example be the case if the pseudo-scalar coupling is mediated by a light particle (*e.g.* a light CP-odd Higgs boson, see below).
If there is helicity-symmetry in the scalar or pseudo-scalar coupling (*i.e.* $C_S = C'_S$ or $C_P = C'_P$), the presence of such contributions from BSM might be not observable in the branching fractions. While there are no other significant constraints on scalar and pseudo-scalar couplings in $b \to s(d)$ transitions than the ones from $B^0_{(s)} \to \mu^+\mu^-$, there are complementary constraints on the axial-vector coupling (C_{10}) from inclusive or exclusive $b \to s\ell^+\ell^-$ decays such as $B^0 \to K^{*0}\mu^+\mu^-$ and $B^\pm \to K^\pm\mu^+\mu^-$. The impact of the combination of these constraints on C_{10} are for example discussed in Ref. [54].

Two-Higgs-Doublet Models

A set of models that can be probed with the decays $B^0_{(s)} \to \mu^+\mu^-$ are models with an extended Higgs sector, including additional scalar and pseudo-scalar Higgs bosons. A simple type of these models are Two-Higgs-Doublet models (2HDM). Besides the Higgs doublet ϕ_1 present in the SM there is an additional Higgs doublet ϕ_2. The 2HDM are categorised into different types according to the couplings of the Higgs doublets: in type-I models only one of the doublets couples to fermions while in type-II models one of the doublets couples to the up-type quarks and the other to the down-type quarks as well as the charged leptons[4].
The second doublet adds four degrees of freedom, which transform into four additional Higgs bosons in the Electroweak Symmetry Breaking. The SM Higgs, h^0, is accompanied by another CP-even Higgs, H^0, a CP-odd Higgs boson, A^0, and two charged Higgs bosons, $H^\pm$. All these additional scalar and pseudo-scalar particles can act as intermediate particles mediating the decays $B^0_{(s)} \to \mu^+\mu^-$ (*cf.* Fig. 6.2) and therefore modify their branching fractions. N.B. that for this a charged Higgs boson is required, which can take part in the flavour changing process.
An important parameter in these models is the ratio $\tan\beta$ between the vacuum expectation values v_1 and v_2 of the two doublets. The branching fractions $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ depend on this parameter as well as on the mass of the charged Higgs, $m_{H^\pm}$, as the scalar and pseudo-scalar couplings $C_{S,P}$ are in the 2HDM-II given by

$$C_S = -C_P \propto \tan^2\beta \frac{\ln(m^2_{H^\pm}/m^2_t)}{1 - m^2_{H^\pm}/m^2_t}, \tag{6.4}$$

where m_t is the top-quark mass [83]. This leads to a $\tan^4\beta$ dependence of the branching fractions in this model. Thus, significantly enhanced branching fractions are expected for large values of $\tan\beta$ (*i.e.* $\tan\beta > 50$) as well as small masses of the charged Higgs bosons (*i.e.* $m_{H^\pm} < 300\,\mathrm{GeV}/c^2$) (*cf.* Fig. 3 in Ref. [193]). On the other hand, destructive interference between the pseudo-scalar and the axial-vector coupling can lead to branching fractions smaller than the SM predictions for values of $m_{H^\pm}$ larger than about $300\,\mathrm{GeV}/c^2$ as well as for $\tan\beta < 25$ [193].

[4] There are also type-III (one doublet couples to quarks, one to the charged leptons) and type-IV models (one doublet couples to the up-type quarks and the charged leptons, the other to the down-type quarks) [124]. These models are not discussed here.

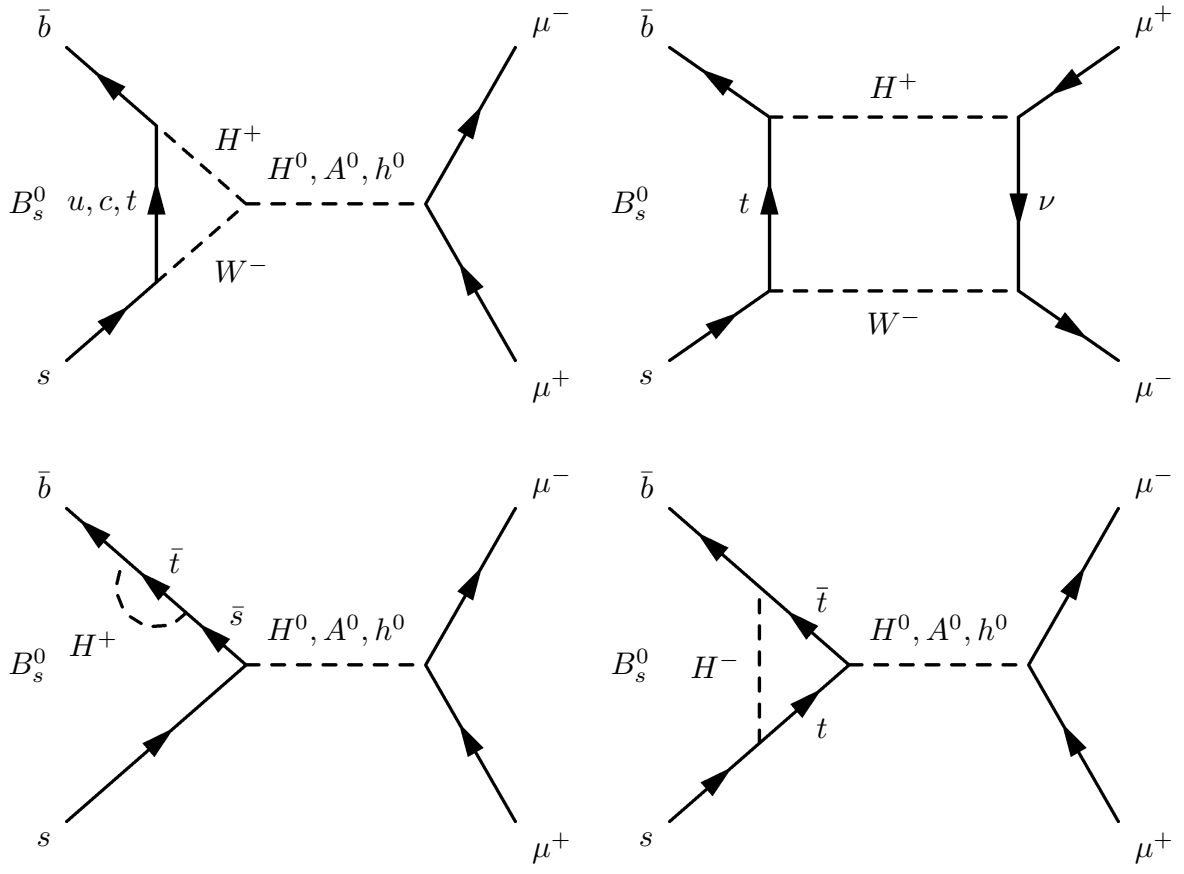


Figure 6.2: Feynman diagrams for the decay $B^0_s \to \mu^+\mu^-$ mediated by charged and neutral Higgs bosons in the 2HDM

Supersymmetric Models

One of the problems of the SM with a Higgs boson at a mass scale around the electroweak scale is that its SM mass is not stabilised. Self-energy corrections (*e.g.* from virtual top-quark loops, *cf.* Fig. 6.3(a)) lead to a quadratic divergence of the Higgs mass. This could be renormalised with a fine-tuning of the Higgs bare mass, but this solution is considered unnatural.

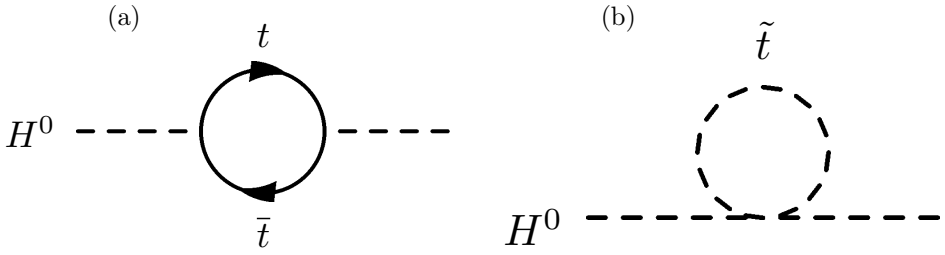


Figure 6.3: Loop corrections of the Higgs mass in (a) SM and in (b) SUSY models

A natural way to compensate for these divergent terms is to introduce an additional symmetry between elementary fermions and bosons. This supersymmetry (SUSY) attributes to each elementary boson in the SM a fermionic superpartner (*e.g.* gluon (g) $\to$ gluino ($\tilde{g}$)) and to each fermion a bosonic superpartner (*e.g.* bottom quark (b) $\to$ sbottom quark ($\tilde{b}$)). The loop corrections to the Higgs mass then also incorporate loop diagrams including the stop (*i.e.* the superpartner of the top) as shown in Fig. 6.3(b). These diagrams act as counter part such that the self-energy correction of the Higgs mass is small as long as the difference between the stop and top mass is not too large.
Besides other phenomena not provided by the SM, there is in many Supersymmetric Models a natural candidate for Dark Matter. A discrete symmetry, called R-parity, with the associated discrete charge

$$R = (-1)^{2s+3B+L}, \tag{6.5}$$

which needs to be conserved in every process[5], can be introduced. As all supersymmetric particles carry odd R-parity and all SM particles even R-parity, the decay of a supersymmetric particle into two SM particles is forbidden. Therefore the lightest supersymmetric particle is stable. If this particle is electrically neutral, such as the neutralino (*i.e.* a linear combinations of the superpartners of the neutral electroweak gauge and Higgs bosons), it could be a good candidate for Cold Dark matter.

The Minimal Supersymmetric Standard Model (MSSM) is one of the most studied SUSY models. It is a supersymmetric extension of the SM with a minimal particle content. It adds – as pointed out above – to each boson in the SM a fermion superpartner and vice versa. Due to the additional symmetry the particle fields must now be described as components of left-handed Weyl spinors. As the superpotential describing the dynamics of the fields must be a holomorphic function, this leads to the need of an additional Higgs doublet to the one present in the SM with similar couplings as in 2HDM.

[5] The parameters B and L are the usual baryon and lepton number. The parameter s is the spin of the particle.

As the supersymmetric partners of the SM particles have not been observed so far, SUSY must be broken at a scale higher than the electroweak scale. There are different mechanisms that could explain this breaking. The three most popular ones for the MSSM are the gravity-mediated, the gauge-mediated and the anomaly-mediated symmetry breaking.
In the first approach, gravity, which can be added at very high energies to the model, is responsible for breaking the symmetry. A minimal (in terms of the particle spectrum) realisation of this option is called minimal Super Gravity (mSUGRA). The other two options use the interaction of the gauge bosons to new messenger fields or loop effects to realise the breaking of SUSY, but both these approaches are disfavoured after the observation of a SM-like Higgs boson at $125\,\mathrm{GeV}/c^2$ [63].
A possible way to include SUSY breaking in the effective Lagrangian of MSSM is to add ad hoc bilinear and trilinear terms of the scalar fields (*Soft SUSY breaking*). The only constraint on those extra terms is that they conserve the cancelation of the quadratic divergences of the self-energy corrections as discussed above.

Constrained Minimal Supersymmetric Standard Model (CMSSM) The CMSSM is the most simple realisation of the MSSM. It assumes that all scalar particles in the model have the same mass m_0 and all gauginos (*i.e.* superpartners of the gauge bosons) the same mass $m_{1/2}$ at the energy scale where the Grand Unification Theory (GUT) is valid. These two masses are two of the additional free parameters in this model. The trilinear coupling A_0, which is also assumed to be the same for all particles at the GUT scale, is another free parameter. The other additional free parameters are the ratio of the vacuum expectation values of the two Higgs doublets, $\tan\beta$, and the sign of the higgsino mass term, μ, which governs the Higgs masses.

Non-Universal Higgs Masses (NUHM) A common relaxation of the assumptions in the CMSSM is that the mass of the Higgs bosons at the GUT scale is not described by m_0, which leads to an additional degree of freedom, m_A. As the Higgs fields and the other scalar fields are part of different multiplets in the Lagrangian, this assumption is very reasonable. This means that μ and not just its sign is a free parameter in this model.

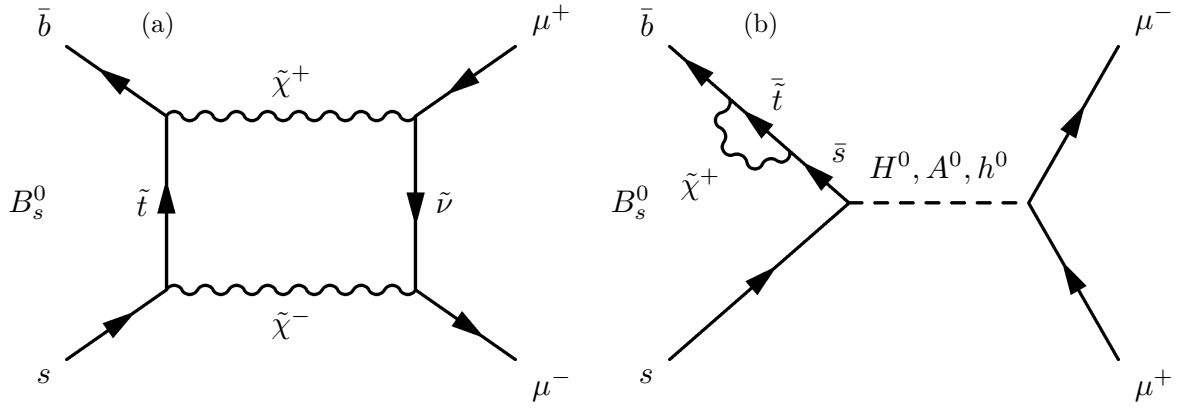


Figure 6.4: (a) Box diagram and (b) Penguin diagram for $B^0_s \to \mu^+\mu^-$ in SUSY including stop quark ($\tilde{t}$), sneutrino ($\tilde{\nu}$) and chargino ($\tilde{\chi}^\pm$) as well as Higgs bosons as intermediate particles.

In the CMSSM as well as in the NUHM the branching fractions of $B^0_{(s)} \to \mu^+\mu^-$ can be significantly modified due to the Higgs bosons and the superpartners of the SM particles as intermediate particles (*cf.* Fig. 6.4). SUSY models can contribute to the branching fractions of the decays $B^0_{(s)} \to \mu^+\mu^-$ not only via new (pseudo-)scalar couplings, but can also lead to modifications of the axial-vector couplings described by $C_{10}^{(\prime)}$ due to the enlarged particle spectrum.
In particular large values of $\tan\beta$ and small masses m_A of the CP-odd Higgs boson can lead to a significant increase in the branching fractions of $B^0_{(s)} \to \mu^+\mu^-$ [95] as the scalar and pseudo-scalar couplings are in these models proportional [134] to

$$C_S = -C_P \propto \frac{\tan^3\beta}{m_A^2}. \tag{6.6}$$

Next-to-Minimal Supersymmetric Standard Model (NMSSM) One of the problems in the CMSSM or NUHM is that due to phenomenological reasons the scale of the higgsino mass term μ needs to be at the electroweak scale rather than at the Planck scale, which would serve as natural cut-off scale [106, 132, 153].
A possible solution is to add a gauge singlet as part of a supermultiplet. This also corresponds to the simplest possible extension of the particle spectrum in the MSSM. The additional degrees of freedom transform into a third CP-even Higgs boson H_1^0, a second CP-odd Higgs boson A_1^0 and an additional fermion called "singlino", which mixes with the neutralinos already present in the MSSM [133].
In this theory the A_1^0 boson might have a mass of the order of $10\,\mathrm{GeV}/c^2$, which would lead to significant destructive interference between the pseudo-scalar and the axial-vector contribution, reducing the branching fractions $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ by a factor three with respect to the SM predictions. On the other hand a light scalar Higgs boson would have no significant impact on the branching fractions [160].

Other Models beyond the SM

Besides the above mentioned models, there are other models beyond the SM, in which new particles can act as additional intermediate particles mediating the decays $B^0_{(s)} \to \mu^+\mu^-$.

Fourth Generation A first example is the extension of the number of quark families, or more general, fermion families to four or more. The additional quarks are usually called top- (t') and bottom-prime quark (b'). The current direct mass limits on such quarks are $m_{b'} > 130\,\mathrm{GeV}/c^2$ and $m_{t'} > 685\,\mathrm{GeV}/c^2$ at 95 % C.L. [6] if they carry the same charge and spin as the SM quarks [81].
The t' quark could act as intermediate particle in $B^0_{(s)} \to \mu^+\mu^-$ in the same manner as the up-type quarks in the SM as shown in Fig. 6.5(a). This would lead to a modification of C_{10} as shown in Eq. (6.3). As the 4×4-counterpart of the CKM-matrix must be unitary and incorporate the present CKM-matrix – close to unitarity – as 3×3-submatrix, the diagrams including t' must suffer large Cabbibo-suppression. Therefore, additional quarks cannot lead to a large enhancement of the branching fractions and especially not simultaneously for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$.

[6] Both limits come from direct searches in pp collisions.

***Z′* and Technicolor Models** Additional gauge bosons such as a high mass counterpart of the Z boson, called Z', could also contribute to the decay as shown in Fig. 6.5(b). In general, vector as well as scalar gauge bosons lead to additional decay amplitudes proportional to $1/M^2_{V,S}$ where $M_{V,S}$ is the mass of these additional bosons.
Limits on the ratio of the Z' coupling, $g_{Z'}$, and its mass, $M_{Z'}$, show that only a Z' boson with a mass smaller than 1 TeV/c^2 can significantly enhance the branching fractions of $B^0_{(s)} \to \mu^+\mu^-$ with respect to their SM predicitions [160].

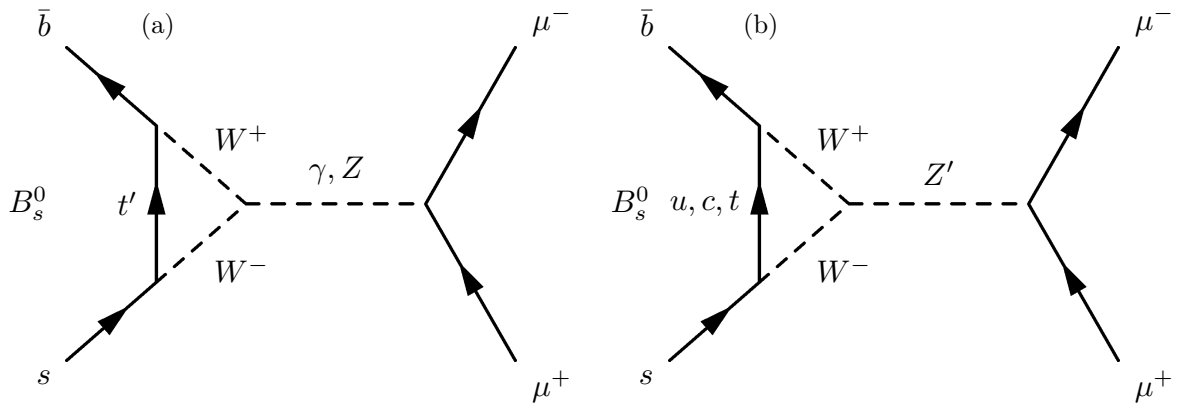


Figure 6.5: Feynman Diagrams for $B^0_s \to \mu^+\mu^-$ in (a) fourth generation model with a t' as intermediate particle and in (b) Z' models

A further example of models, in which the branching fractions of $B^0_{(s)} \to \mu^+\mu^-$ can be significantly modified, are Technicolor Models [192, 209]. The basic idea of these models is that the breaking of the electroweak symmetry is caused by a new gauge interaction coupling to fermions, which behaves similarly as the strong interaction: at very high energies it is asymptotically free while there is confinement at energies around the electroweak scale. This additional interaction spontaneously breaks the chiral symmetry of the massless fermions and therefore the electroweak gauge symmetry giving the W and Z bosons their masses.
This mechanism leads to additional particles such as a Z' boson and – depending on the detailed realisation of the model – heavy scalar particles, which can act as intermediate particles in the decays $B^0_{(s)} \to \mu^+\mu^-$.
Due to the discovery of a SM-like Higgs boson, Technicolor Models explaining the electroweak symmetry breaking are considered obsolete, although they do not require a fine-tuning of the Higgs mass as discussed above.

Minimal Flavour Violation

A general concept for BSM models is the hypothesis of Minimal Flavour Violation (MFV). MFV assumes that the flavour structure in the considered model describing physics beyond the SM is not different from the structure present in the SM. Therefore the only source of flavour-symmetry breaking are the Yukawa couplings discussed in the first chapter, and the CKM matrix is still governing the flavour mixing. Obviously, the MFV hypothesis does not hold for models with an extended flavour structure (*e.g.* fourth generation models with flavour changing interactions

between the fourth and the other generations).
An especially good observable to test this hypothesis is the ratio of the branching fractions $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ with the SM prediction

$$\frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B^0_s \to \mu^+\mu^-)} = \frac{m_{B^0}\tau_{B^0} f^2_{B^0}}{m_{B^0_s}\tau_{B^0_s} f^2_{B^0_s}} \cdot \left|\frac{V_{td}}{V_{ts}}\right|^2. \tag{6.7}$$

This observable has the advantage that several uncertainties – both in the theoretical calculation as well as in the experimental measurement – cancel in the ratio. The SM prediction [97] is

$$\frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B^0_s \to \mu^+\mu^-)} = 0.0329 \times \left(1 \pm 0.06 \pm 2\sigma^r_{f_{s/d}}\right), \tag{6.8}$$

where the decay constant ratio $f_{B^0}/f_{B^0_s} = 1.195$ is used with its relative uncertainty $\sigma^r_{f_{s/d}}$. The first quoted uncertainty is due to the other theoretical uncertainties. Any significant deviation from this prediction would be a clear sign of the violation of the MFV hypothesis and thus of the presence of physics beyond the SM.

6.2 Pre-LHC Experimental Status

Several experiments studied before the start of the LHC the branching fractions of $B^0_s \to \mu^+\mu^-$ (UA1 [49], L3 [42] and D0 [40]) or $B^0 \to \mu^+\mu^-$ (ARGUS [50], UA1 [49], BELLE [107], BaBar [67], CLEO [80] and L3 [42]).
The most stringent upper limits on the two branching fractions came from the CDF experiment at the TeVatron [3] with

$$\begin{aligned} \mathcal{B}(B^0_s \to \mu^+\mu^-) &< 4.3 \times 10^{-8} \text{ @ } 95\,\% \text{ C.L.} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &< 7.6 \times 10^{-9} \text{ @ } 95\,\% \text{ C.L.} \end{aligned}$$

These upper limits – in both decays more than one order of magnitude larger than the SM prediction – still kept the possibility of a significant deviation from the SM prediction, allowing for example large values of $\tan\beta$ in MSSM and 2HDM.

6.3 The Search Strategy for $B^0_{(s)} \to \mu^+\mu^-$ at LHCb

The LHCb experiment has performed a search for $B^0_{(s)} \to \mu^+\mu^-$ based on the full 2011 (1 fb^{-1} at $\sqrt{s} = 7\,\text{TeV}$) and 2012 (2 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$) data samples [25] following a similar analysis strategy as in its previous searches [19, 21, 22, 24]. The strategy of this analysis can be split into five steps:

1. **Pre-selection and blinding** The goal of the pre-selection stage is to reduce the data sample to a manageable size while keeping the signal efficiency as high as possible. The signal candidates for the search are created from two well-reconstructed and -identified oppositely charged muons that are detached from any primary vertex in the event[7]. Details of the

[7] The selection has been done for 2011 and 2012 data with DaVinci v33r1 using the database tags `cond-20121116` and `dddb-20120831`.

selection criteria can be found in Tab. 6.1. The invariant mass of the dimuon combination must be within [4.9,6.0] GeV/c^2. The momentum vector of the $B^0_{(s)}$ candidate has to point back towards its associated primary vertex defined by the one, to which it has the smallest impact parameter. It needs to have a transverse momentum with respect to the beam axis (p_{T}) larger than 0.5 GeV/c^2, and its decay time should not exceed nine times the nominal B^0_s lifetime.

Table 6.1: Selection cuts for $B^0_{(s)} \to \mu^+\mu^-$ candidates as well as for candidates in the control and normalisation channels $B^0_{(s)} \to h^+h'^-$, $B^\pm \to J/\psi\, K^\pm$ and $B^0_s \to J/\psi\, \phi$.

Variable	$\boldsymbol{B^0_{(s)} \to \mu^+\mu^-}$ **and** $\boldsymbol{B^0_{(s)} \to h^+h'^-}$	$\boldsymbol{B^\pm \to J/\psi K^\pm}$ **and** $\boldsymbol{B^0_s \to J/\psi\phi}$
Muons/Hadrons		
χ^2_{track}/ndf	< 3	< 3
Ghost prob.[a]	< 0.3	–
DOCA[b]	< 0.3 mm	< 0.3 mm
χ^2_{IP}	> 25	> 25
p_{T}	> 250 MeV/c and < 40 GeV/c	> 250 MeV/c and < 40 GeV/c
p	< 500 GeV/c	< 500 GeV/c
$\Delta\mathrm{LL}_{K-\pi}$ (only muons)	< 10	–
$\Delta\mathrm{LL}_{\mu-\pi}$ (only muons)	> -5	–
IsMuon[c] (only muons)	yes	yes
Dimuon/Dihadron		
χ^2_{vtx}/ndf	< 9	< 9
FDS[d]	> 15	> 15
***B* candidate**		
χ^2_{IP}	< 25	< 25
t	$< 9 \cdot \tau_{B^0_s}$	$< 9 \cdot \tau_{B^0_s}$
p_{T}	> 500 MeV/c	–

[a] The ghost probability is a multivariate classifier with values between 0 and 1 to distinguish real tracks from ghost tracks (*i.e.* combination of unrelated hits or track segments in the tracking stations) using kinematical variables as well as the difference between observed and expected hits in the tracking sub-detectors as input. The classifier is built in such a manner that its distribution is flat for ghost tracks.

[b] Distance of closest approach between the two muons

[c] Muon identification as described in Sec. 3.4.1

[d] Flight distance significance of the dimuon/dihadron combination (*i.e.* distance between the production and the decay vertex divided by its uncertainty)

Besides this cut based selection there is a boosted-decision tree (BDT) used to further reduce the background. The six variables entering this classifier are

- the fit quality of the dimuon vertex (χ^2_{vtx});
- the impact parameter (IP(B)) and impact parameter χ^2 ($\chi^2_{\mathrm{IP}}(B)$) of the $B^0_{(s)}$ candi-

date with respect to the primary vertex (PV)[8];

- the minimum distance between the two daughter tracks (DOCA);
- the smaller of the two impact parameters of the daughter tracks with respect to any primary vertex ($\min \mathrm{IP}(\mu)$);
- and the angle between the direction of the $B^0_{(s)}$ candidate momentum and the direction defined by the primary and the dimuon vertices (DIRA).

The classifier is trained on Monte Carlo (MC) using a sample of $B^0_s \to \mu^+\mu^-$ candidates for signal and a sample of $b\bar{b} \to \mu^+\mu^- X$ candidates for background. This variable will be referred to as BDTS hereafter.
A suitable choice as cut on the BDTS response between 0 and 1 has been found to be BDTS > 0.05. This allows to remove about 75 % of the background while keeping 95 % of the signal.
To unbias the analysis from any signal information, a region of $\pm 60\,\mathrm{MeV}/c^2$ in the invariant dimuon mass around the masses of the nominal B^0 and B^0_s mesons is blinded in data. Candidates within these windows are only looked at in the last stage of the analysis, after all selection criteria and normalisations as well as calibrations have been fixed.

Besides the signal sample, several calibration, normalisation and control samples are selected. These samples consist of candidates of the decays $B^0_{(s)} \to h^+h'^-$ (where $h^{(\prime)}$ is either a kaon or a pion), $B^\pm \to J/\psi\, K^\pm$ (where the J/ψ decays into two muons) and $B^0_s \to J/\psi\, \phi$ (where the J/ψ decays also into two muons and the ϕ into two oppositely charged kaons). Their selection criteria are also listed in Tab. 6.1. For $B^0_{(s)} \to h^+h'^-$ the criteria are the same as for the signal channel except the particle identification (PID) requirements. In the case of $B^\pm \to J/\psi\, K^\pm$ and $B^0_s \to J/\psi\, \phi$ the BDTS is defined using the two muon tracks to calculate χ^2_{vtx}, $\min \mathrm{IP}(\mu)$ and DOCA.
The trigger selection for the channels with muons relies on the single and dimuon triggers at the L0. These decisions are refined by single and dimuon trigger lines in the HLT1. The $B^0_{(s)} \to \mu^+\mu^-$ channels are triggered in the HLT2 by a dedicated trigger line for dimuon candidates having a mass compatible with the mass of a b hadron. The channels containing a J/ψ are triggered by a trigger line selecting detached J/ψ candidates. A small fraction of candidates in the signal channels as well as in the channels containing a J/ψ are selected by topological lines designed to select partially reconstructed B decays.
These topological lines are also the main trigger lines for the selection of the $B^0_{(s)} \to h^+h'^-$ candidates. At the lower stages the $B^0_{(s)} \to h^+h'^-$ candidates are selected by the L0 Hadron trigger and trigger lines selecting events with high-p_{T} tracks in the HLT1. More details about the trigger in LHCb can be found in Sec. 3.5.

2. **Signal and background discrimination** The signal and background candidates are separated using another BDT as multivariate classifier based on topological and kinematical properties of $B^0_{(s)} \to \mu^+\mu^-$. Its definition is given in Chapter 7 where also the calibration of the BDT distribution for signal is described. For background the calibration is done with the invariant mass sidebands of the signal sample. The BDT response

[8] The impact parameter χ^2 is defined as the difference in the fit χ^2 of the PV with and without the $B^0_{(s)}$ candidate. It behaves almost as the square of the IP divided by its uncertainty.

can range from 0 (background-like) to 1 (signal-like), and its distribution should be uniform for signal per construction. The BDT is evaluated in eight bins with boundaries $[0.00, 0.25, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00]$
The invariant dimuon mass $m_{\mu^+\mu^-}$ is also used to discriminate signal and background. The estimation of its distribution for signal is described in Chapter 8 while for background the calibration is also done based on the sidebands of the signal sample using an exponential distribution to describe the shape.

3. **Normalisation of the branching fraction** The analysis does not use the absolute luminosity scale to translate the number of signal candidates to a branching fraction value, but uses instead other B decays with a well-known branching fraction. The two used decay channels are $B^0 \to K^+\pi^-$ and $B^\pm \to J/\psi K^\pm$. While the former channel is very similar to the signal in terms of the event topology, the latter is similar in terms of trigger and PID to the signal channels due to the two final state muons.
For both signal channels, the ratio of the signal branching fraction $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ to the signal yield $N_{B^0_{(s)}\to\mu^+\mu^-}$, $\alpha_{B^0_{(s)}\to\mu^+\mu^-}$, is determined as

$$\begin{aligned} \alpha_{B^0_{(s)}\to\mu^+\mu^-} &= \frac{\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)}{N_{B^0_{(s)}\to\mu^+\mu^-}} \\ &= \frac{f_{d/s}}{f_d} \cdot \frac{\varepsilon^{\text{TRIG|SEL}}_{\text{sig}}}{\varepsilon^{\text{TRIG|SEL}}_{\text{calib}}} \frac{\varepsilon^{\text{SEL|REC}}_{\text{sig}}}{\varepsilon^{\text{SEL|REC}}_{\text{calib}}} \frac{\varepsilon^{\text{REC}}_{\text{sig}}}{\varepsilon^{\text{REC}}_{\text{calib}}} \cdot \frac{\mathcal{B}_{\text{calib}}}{N_{\text{calib}}}, \end{aligned} \tag{6.9}$$

where $f_s/f_d = 0.259 \pm 0.015$ (statistical and systematic uncertainties combined) is the hadronisation ratio between B^0_s and $B^\pm/B^0$ mesons [26]. Further, N_{calib} is the signal yield of the calibration channel, and $\mathcal{B}_{\text{calib}}$ the known branching fraction of the calibration channel; $\varepsilon^{\text{TRIG|SEL}}$ is the trigger efficiency for selected candidates, $\varepsilon^{\text{SEL|REC}}$ the selection efficiency for reconstructed candidates and ε^{REC} the reconstruction efficiency where the subscript indicates the signal or the calibration channel in Eq. (6.9).
The obtained values for $\alpha_{B^0_s\to\mu^+\mu^-}$ and $\alpha_{B^0\to\mu^+\mu^-}$ averaged for 2011 and 2012 data are

$$\alpha_{B^0_s\to\mu^+\mu^-} = (8.93 \pm 0.64) \times 10^{-11}$$

and

$$\alpha_{B^0\to\mu^+\mu^-} = (2.38 \pm 0.09) \times 10^{-11}.$$

These numbers can be translated into expected signal yields in the full 2011 and 2012 data samples for the two channels under the SM hypothesis, which are $N_{B^0_s\to\mu^+\mu^-} = 39 \pm 4$ and $N_{B^0\to\mu^+\mu^-} = 4.5 \pm 0.4$ [25].

4. **Estimation of Exclusive Background Channels** The main source of background is $b\bar{b} \to \mu^+\mu^- X$. But, there are other channels, which can fake a signal. These are separated into three categories:

   - **Double mis-ID:** the decays $B^0_{(s)} \to h^+h'^-$ where both final state particles are misidentified as muons, either by the detector or due to decays in flight;

- **Semileptonic decays:** the decays $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_s \to K^-\mu^+\nu_\mu$ and $\Lambda^0_b \to p\mu^-\overline{\nu}_\mu$ where the charged non-muonic final state particle is mis-identified as a muon;
- **Partially reconstructed decays:** decays like $B^{0,\pm} \to \mu^+\mu^-\pi^{0,\pm}$ where there exists a true dimuon vertex, but at least one other particle that is missed.

The $m_{\mu^+\mu^-}$ and BDT distributions of these decays are estimated from MC while their absolute yields in the data sample are estimated using the channels $B^0_{(s)} \to h^+h'^-$ – for the first category – and $B^\pm \to J/\psi\, K^\pm$ as normalisation channels.
Overall only the decays $B^0_{(s)} \to h^+h'^-$, $B^0 \to \pi^-\mu^+\nu_\mu$ and $B^{0,\pm} \to \mu^+\mu^-\pi^{0,\pm}$ have been found to be relevant, *i.e.* leading to a significant yield in the most sensitive BDT region and having an invariant dimuon mass distribution different from the one for $b\bar{b} \to \mu^+\mu^- X$. These channels are later included in the branching fraction fit as well as in the calculation of upper limits on the branching fractions.

5. **Results** The results are either extracted with a branching fraction fit or – if the result is not significant – with a binned $\mathrm{CL_s}$ method [210] to get an upper limit on the branching fraction.
The branching fraction fit is performed as an unbinned maximum-likelihood fit of the invariant dimuon mass in the range 4.9 to 6.0 GeV/c^2 simultaneously in the 8 BDT bins. The distributions in the dimuon mass and the BDT are for the signal channels as well as the exclusive background channels fixed to those determined in the steps described above. The uncertainties on the parameters describing distributions are treated as Gaussian constraints. Furthermore, also the total yields of the exclusive background channels are fixed with their uncertainties as Gaussian constraints. Only the shape and the yield of the combinatorial background in the eight BDT bins described by exponential distributions and the branching fractions of $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are treated as free parameters.

7

The Calibration of the Multivariate Classifier for Signal

The analysis of $B^0_{(s)} \to \mu^+\mu^-$ uses a boosted-decision tree (BDT) [89,211] to separate signal and background. Background decays are mostly due to events with two semileptonic decays of a $b\bar{b}$ pair ($b\bar{b} \to \mu^+\mu^- X$; *cf.* Fig. 7.1). The purpose of this multivariate classifier is to select two-body decays of a b hadron based on topological and kinematical variables. The BDT is trained with Monte Carlo samples for signal as well as background. To avoid biases induced by differences between simulation and data, the trained BDT response to both signal and background is calibrated on data. For background, the invariant mass sidebands of the signal candidates are used while the decays $B^0_{(s)} \to h^+h'^-$ with h' being a charged kaon or pion are used as a signal proxy. This chapter describes the details of the calibration procedure. Section 7.1 describes the BDT, while the simulation studies for the calibration are discussed in Sec. 7.2. Section 7.3 addresses the determination and evaluation of the efficiency to distinguish kaons and pions. The actual calibration is described in Sec. 7.4.

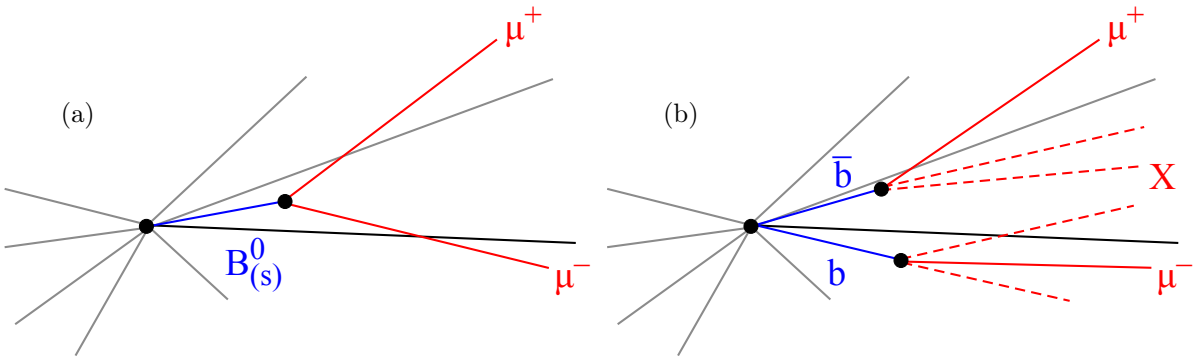


Figure 7.1: The topologies of the decays (a) $B^0_{(s)} \to \mu^+\mu^-$ and (b) $b\bar{b} \to \mu^+\mu^- X$

7.1 The Boosted Decision Tree

The boosted decision tree (BDT) uses in total twelve variables:

- the $B^0_{(s)}$ candidate decay time (t);
- the smaller of the two impact parameter significance values of the muons (IPS(μ));
- the impact parameter of the $B^0_{(s)}$ candidate (IP(B));
- the distance of closest approach between the two muons (DOCA);
- the isolation of the two muons with respect to any other track in the event[1] (I(μ));
- the transverse momentum of the $B^0_{(s)}$ candidate ($p_{\mathrm{T}}(B)$);
- the cosine of the angle between the muon momentum in the dimuon rest frame and the vector perpendicular to the beam axis and the momentum of the $B^0_{(s)}$ candidate ($\cos P$);
- the isolation of the $B^0_{(s)}$ candidate, defined as $p_{\mathrm{T}}(B^0_{(s)})/(p_{\mathrm{T}}(B^0_{(s)}) + \sum_i p_{\mathrm{T}i})$ where the sum runs over all tracks other than the two muon candidates in a cone around the $B^0_{(s)}$ candidate with a radius of $R = \sqrt{\Delta\phi_i^2 + \Delta\eta_i^2} = 1.0$ [2] (I(B));
- the angle between the momentum of the $B^0_{(s)}$ candidate and the thrust momentum of the $B^0_{(s)}$, defined as the sum of the momenta of all long tracks originating from the same primary vertex as the $B^0_{(s)}$ candidate excluding those due to decays of long lived particles. If no such track is found, the variable is set to zero (other B angle);
- the angle between the direction of the positive muon candidate in the rest frame of the $B^0_{(s)}$ and the thrust momentum in the $B^0_{(s)}$ rest frame (B boost);
- the absolute value of the difference between the pseudorapidity of the two muon candidates ($|\Delta\eta|$);
- the absolute value of the difference between the azimuthal angle ϕ of the two muon candidates ($|\Delta\phi|$).

The distributions of the variables for signal and background are shown in Figs 7.2 and 7.3. The BDT is trained using an MC sample of two million $B^0_s \to \mu^+\mu^-$ events for signal and an MC sample of about 120 million $b\bar{b} \to \mu^+\mu^- X$ events for background. The signal sample was created in the MC 2012 production[3], while about 100 million events of the background sample were created in the MC 10 production[4] and the rest was also produced in the MC 2012 production.

[1] The variable counts the number of detached tracks (impact parameter significance larger than 3) that form with at least one of the muon candidates a good vertex.

[2] $\Delta\phi_i$ and $\Delta\eta_i$ represent the separation between track i and the $B^0_{(s)}$ candidate in the azimuthal angle and the pseudorapidity, respectively.

[3] The production has been using the Sim06b conditions and the reconstruction version Reco14. The selection is based on DaVinci v33r1 using the same database tag as for signal (`cond-20121116` and `dddb-20120831`).

[4] The production has been using the Sim01 conditions and the reconstruction version Reco08. The selection is based on DaVinci v33r1 using the same database tag as for signal (`cond-20121116` and `dddb-20120831`).

The BDT is trained using the AdaBoost algorithm [214]. After training the BDT is transformed such that its output is in the range between zero and one and should follow a flat distribution for signal.

7.2 Monte Carlo Studies for the MV Classifier Calibration

The BDT calibration for signal uses information from simulated $B_s^0 \to \mu^+\mu^-$, $B^0 \to K^\pm\pi^\mp$ and $\Lambda_b^0 \to ph^-$ events to obtain an appropriate description of the invariant mass distribution shape for $B_{(s)}^0 \to h^+h'^-$.

7.2.1 $B_s^0 \to \mu^+\mu^-$ MC Study

The used sample of simulated $B_s^0 \to \mu^+\mu^-$ events is the same as that used for the training of the BDT and has a size of half a million events after applying the same selection as for the $B_{(s)}^0 \to \mu^+\mu^-$ data sample (*cf.* Tab. 6.1) and requiring an MC-truth matching of the B_s^0 meson. The invariant dimuon mass ($m_{\mu^+\mu^-}$) distribution is shown in Fig. 7.4 and fitted by the sum of two Crystal Ball functions (${}_d\mathcal{CB}$) with tails to opposite sides and different tail parameters, but with the same central value μ and width σ of the Gaussian kernel of the Crystal Ball functions

$$\begin{aligned} {}_d\mathcal{CB} = {} & f \cdot \mathcal{CB}(m_{\mu^+\mu^-}|\mu,\sigma,\alpha_l,n_l) \\ & + (1-f)\cdot \mathcal{CB}(m_{\mu^+\mu^-}|\mu,\sigma,\alpha_r,n_r) \end{aligned} \tag{7.1}$$

with the Crystal Ball function ($\mathcal{CB}$) [218]

$$\mathcal{CB}(m_{\mu^+\mu^-}|\mu,\sigma,\alpha,n) = N \cdot \begin{cases} e^{-\frac{(m_{\mu^+\mu^-}-\mu)^2}{2\sigma^2}}, & \text{if } \frac{m_{\mu^+\mu^-}-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{\alpha^2}{2}} \left(\frac{n}{|\alpha|} - |\alpha| - \frac{m_{\mu^+\mu^-}-\mu}{\sigma}\right)^{-n} & \text{else} \end{cases} \tag{7.2}$$

$$\text{with } N = \left[\sigma\left(\frac{n}{|\alpha|}\frac{1}{n-1}e^{-\frac{\alpha^2}{2}} + \sqrt{\frac{\pi}{2}}\left[1+\operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right]\right)\right]^{-1},$$

where $f \in [0,1]$ describes the fraction of the first Crystal Ball function and n_l, n_r, $\alpha_l > 0$ and $\alpha_r < 0$ describe the exponential tails to the left (l) and right (r) side and erf is the error function. The fit is performed by a binned maximum-likelihood algorithm. The resulting parameters are listed in Tab. 7.1.

7.2.2 $B^0 \to K^\pm\pi^\mp$ MC Study

Also the MC sample for $B^0 \to K^\pm\pi^\mp$ has been generated in the MC 2012 production. It contains about 125'000 events after applying the same selection as for the $B_s^0 \to \mu^+\mu^-$ data sample (*cf.* Tab. 6.1) and requiring an MC-truth matching of the B^0 meson.
Also the invariant mass distribution for the pion-kaon combination, $m_{K\pi}$, is described by a double Crystal Ball function matched to the MC sample in a binned maximum-likelihood fit (*cf.* Fig. 7.5). The fitted parameters can be found in Tab. 7.1. The tail parameters $n_{r,l}$ and $\alpha_{r,l}$ describing the non-Gaussian part of the invariant mass resolution are different than in the $B_s^0 \to \mu^+\mu^-$ MC sample as the hadronic particles interact differently with the detector material than muons.

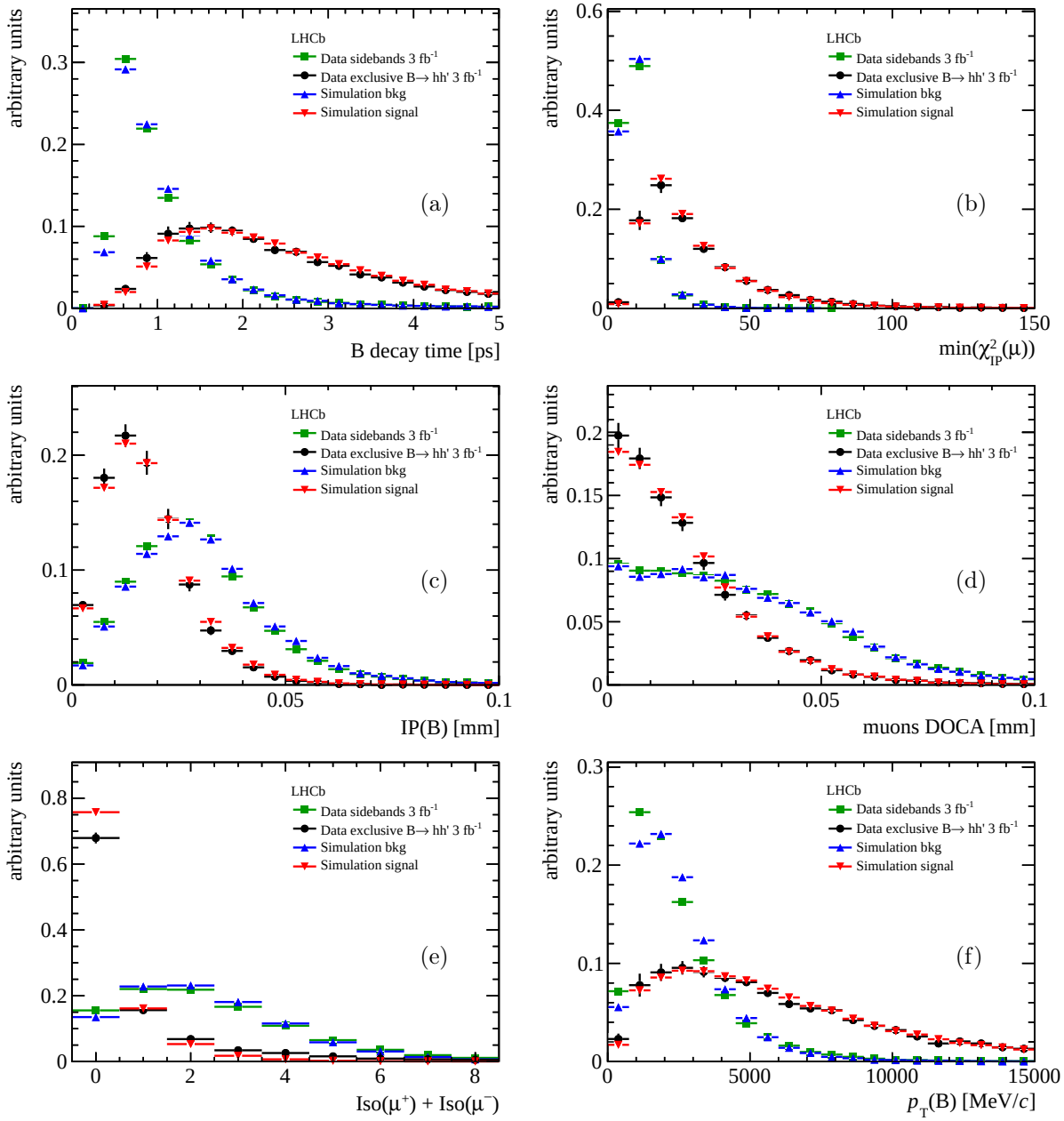


Figure 7.2: BDT input variables (part 1), the distributions for $B_s^0 \to \mu^+\mu^-$ (red downward pointing triangles) and $b\bar{b} \to \mu^+\mu^- X$ (blue upward pointing triangles) in MC as well as for the $B_s^0 \to \mu^+\mu^-$ data sidebands (green squares) and background subtracted $B_{(s)}^0 \to h^+h'^-$ data (black dots) are shown: (a) $B_{(s)}^0$ candidate proper time; (b) minimum impact parameter significance of the muons; (c) impact parameter of the $B_{(s)}^0$ candidate; (d) distance of closest approach between the two muons; (e) isolation of the two muons with respect to any other track in the event; (f) transverse momentum of the $B_{(s)}^0$ candidate

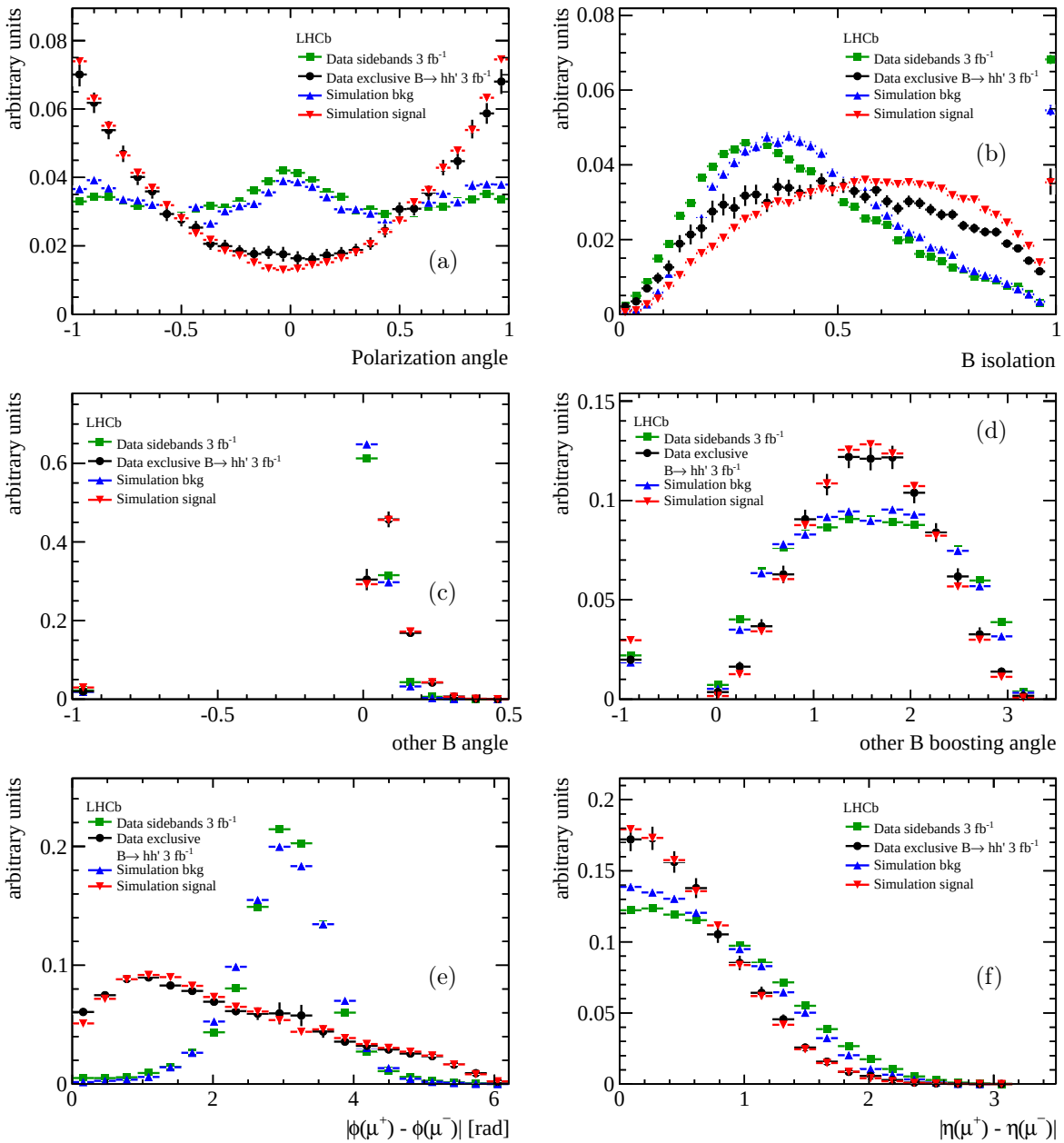


Figure 7.3: BDT input variables (part 2), the distributions for $B^0_s \to \mu^+\mu^-$ (red downward pointing triangles) and $b\bar{b} \to \mu^+\mu^- X$ (blue upward pointing triangles) in MC as well as for the $B^0_s \to \mu^+\mu^-$ data sidebands (green squares) and background subtracted $B^0_{(s)} \to h^+h'^-$ data (black dots) are shown: (a) cosine of the polarisation angle; (b) isolation of the $B^0_{(s)}$ candidate; (c) angle between the momentum of the $B^0_{(s)}$ candidate and the thrust axis; (d) angle between the momentum of the positive muon and the thrust axis; (e) absolute difference in ϕ between the two muons; (f) absolute difference in η between the two muons

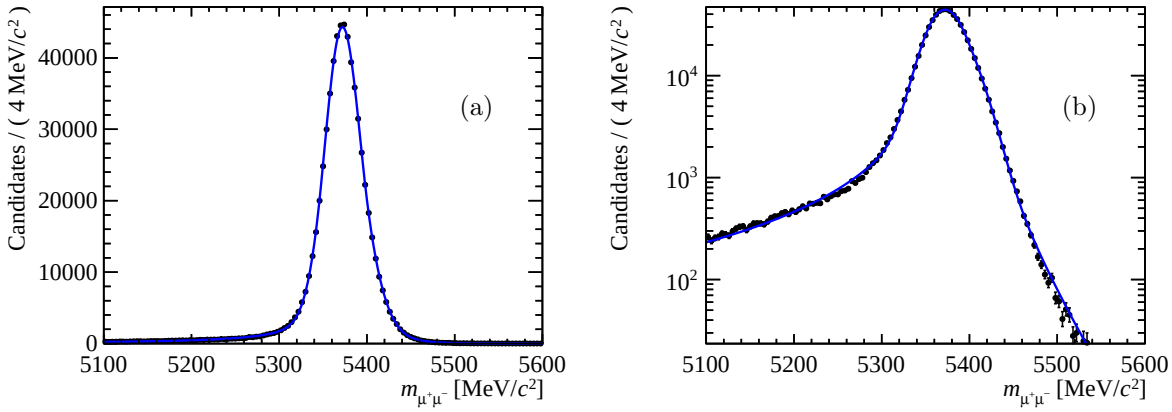


Figure 7.4: The invariant dimuon mass distribution of $B_s^0 \to \mu^+\mu^-$ in MC on (a) linear and (b) log scale. The blue line shows the fitted sum of two Crystal Ball functions.

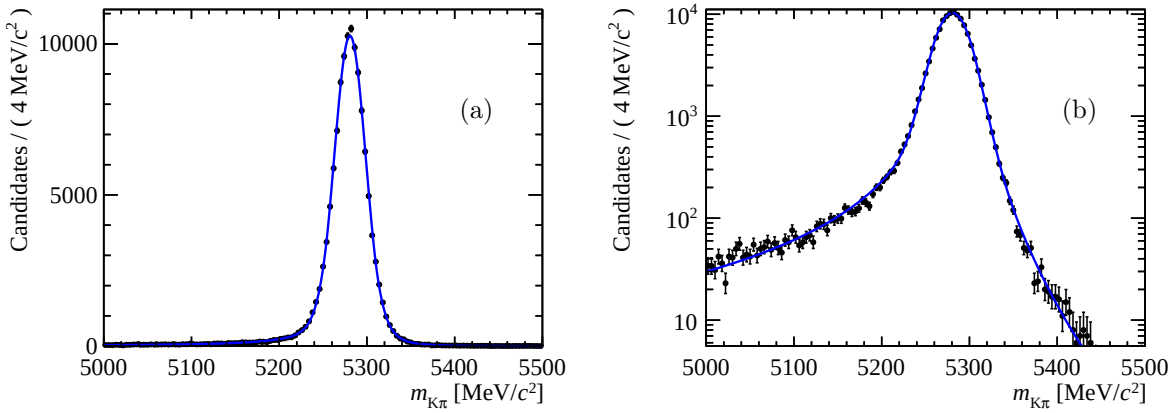


Figure 7.5: The invariant kaon-pion mass distribution of $B^0 \to K^\pm\pi^\mp$ in MC on (a) linear and (b) log scale. The blue line shows the fitted sum of two Crystal Ball functions.

7.2.3 $\Lambda_b^0 \to ph^-$ MC Study

The decays $\Lambda_b^0 \to ph^-$, where h is either a pion or a kaon, are a background to the decays $B^0_{(s)} \to h^+h'^-$ if the proton is mis-identified as a pion or kaon. Thus, the invariant mass distribution of proton-pion/kaon combinations in the decays $\Lambda_b^0 \to p\pi^-$ and $\Lambda_b^0 \to pK^-$ is studied, where the proton is mis-identified as a pion or a kaon. The study is performed on MC samples for $\Lambda_b^0 \to p\pi^-$ and $\Lambda_b^0 \to pK^-$ with two million events each (one million per magnet polarity).
The candidates from the two channels are weighted according to the branching fractions ($\mathcal{B}(\Lambda_b^0 \to p\pi^-) = (4.0 \pm 0.8) \times 10^{-6}$, $\mathcal{B}(\Lambda_b^0 \to pK^-) = (4.8 \pm 0.9) \times 10^{-6}$ [81]) when building the invariant mass distribution. A mis-identification of the proton as a pion or a kaon is forced according to the reconstructed $\Delta\mathrm{LL}_{K-\pi}$ value of the proton[5]: if it is larger than zero, the proton is treated as a kaon, otherwise as a pion.
A double Crystal Ball function is used to describe the resulting distribution, where the width of the two components (σ_l and σ_r) are not common while the central value μ is common for both components.
The obtained invariant dihadron mass distribution together with the binned maximum-likelihood fit is displayed in Fig. 7.6 and the values of the fitted parameters are listed in Tab. 7.1.
The chosen functional form describes the distribution very well. Tests have been performed to check the impact of the BDT selection on the distribution as well as of the fact that the performance of the $\Delta\mathrm{LL}_{K-\pi}$ variable is different between MC and data. Both tests have shown no significant systematic effect on the results of the fit.

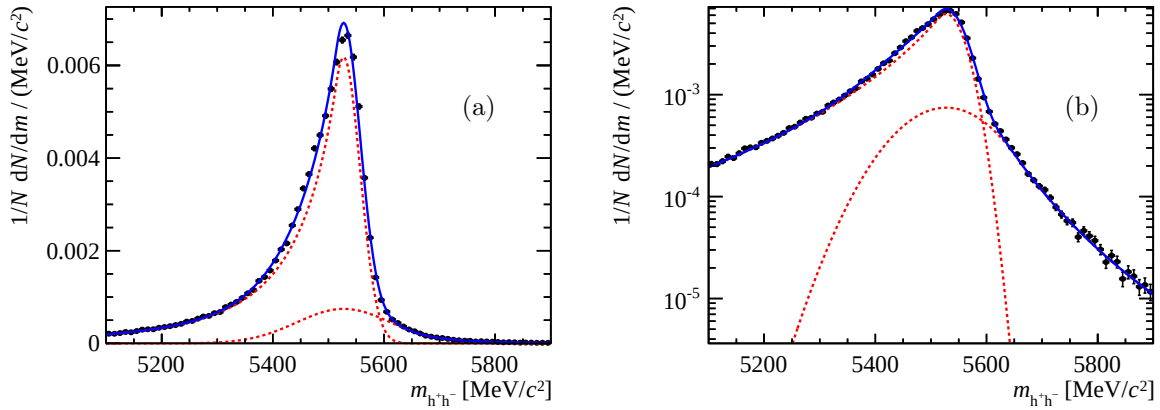


Figure 7.6: The invariant dihadron mass distribution of $\Lambda_b^0 \to ph^-$, where the proton is mis-identified as a pion or a kaon, in MC on (a) linear and (b) log scale. The blue line shows the sum of two Crystal Ball functions with tails to opposite sides and different width of their Gaussian kernels, but common central value. The individual Crystal Ball functions are displayed by the red dashed lines.

7.3 Determination and Evaluation of the PID efficiency

Particle identification information is used in the construction of the invariant dihadron mass distribution $m_{h^+h'^-}$ for the BDT calibration to separate the different decay channels $B^0_{(s)} \to$

[5] The $\Delta\mathrm{LL}_{K-\pi}$ classifier used for particle identification is discussed in detail in the next section.

Table 7.1: Mass parameters for the invariant mass distributions of $B_s^0 \to \mu^+\mu^-$, $B^0 \to K^\pm\pi^\mp$ and $\Lambda_b^0 \to ph^-$ (where the proton is mis-identified as a kaon or a pion) in Monte Carlo extracted from maximum-likelihood fits. The quoted uncertainties those returned by the fit.

Parameter	$\boldsymbol{B_s^0 \to \mu^+\mu^-}$	$\boldsymbol{B^0 \to K^\pm\pi^\mp}$	$\boldsymbol{\Lambda_b^0 \to ph^-}$
μ [MeV/c^2]	5371.842±0.005	5279.930±0.035	5530.985±0.003
σ [MeV/c^2]	20.907±0.008	20.612±0.040	–
σ_l [MeV/c^2]	–	–	26.334±0.004
σ_r [MeV/c^2]	–	–	66 ±7
n_l	1.326±0.013	1.612±0.024	6.041±0.019
n_r	10.560±0.020	4.236±0.030	7.639±0.008
α_l	1.326±0.013	1.612±0.024	0.254±0.015
α_r	-1.302±0.019	-1.558±0.025	-1.164±0.023
f	0.418±0.010	0.448±0.010	0.714±0.014

$\pi^+\pi^-$, $B^0_{(s)} \to K^+\pi^-$, $B^0_{(s)} \to \pi^+K^-$ and $B^0_{(s)} \to K^+K^-$, so to assign each candidate the proper mass hypothesis of the final state particles.
This information is in LHCb described by a delta log-likelihood ΔLL_{x-y} based on information from different sub-detectors (ECAL, HCAL, Muon System, RICH detectors). The higher its value is, the more likely the particle is of type x and not of type y. A value of zero should by construction show that the two hypothesis x and y are equally likely, but the performance of the variable in data and simulation deviates from this assumption.
The delta log-likelihood $\Delta\text{LL}_{K-\pi}$ is used in the calibration to separate the different channels in $B^0_{(s)} \to h^+h'^-$.

7.3.1 Estimation of the PID efficiency

The efficiency for cuts on $\Delta\text{LL}_{K-\pi}$ is determined from samples of $D^{*+} \to D^0\pi^+$ ($D^{*-} \to \overline{D}^0\pi^-$) candidates with $D^0 \to K^-\pi^+$ ($\overline{D}^0 \to K^+\pi^-$). The charge of the D^0 decay products can be identified by the charge of the slow pion from the D^{*+} decay as D^0 decays to $K^-\pi^+$ while the decay to $K^+\pi^-$ is doubly Cabibbo-suppressed. Residual background contributions from the Cabibbo-suppressed decays $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ are corrected for by applying the ${}_s\mathcal{P}lot$ method [208] based on the invariant mass of the D^0 candidate and the mass difference δm between the $D^{*\pm}$ and D^0 candidates.
As shown in Fig. 7.7, the performance of the $\Delta\text{LL}_{K-\pi}$ strongly depends on the track momentum, p, the pseudorapidity of the track, η, and the number of tracks in the event. Therefore the efficiency is determined as a function of p, η and number of tracks as well as separately for the two magnet polarities, called MagDown and MagUp.
The tools to calculate the efficiencies (PIDCalibTools) are provided by the PID group in LHCb [28] and have been used to create dedicated efficiency calculations for the $B^0_{(s)} \to \mu^+\mu^-$ analysis. The efficiencies are estimated in 625 three-dimensional bins (25 momentum bins and five bins each in η and number of tracks) of the parameter space defined as

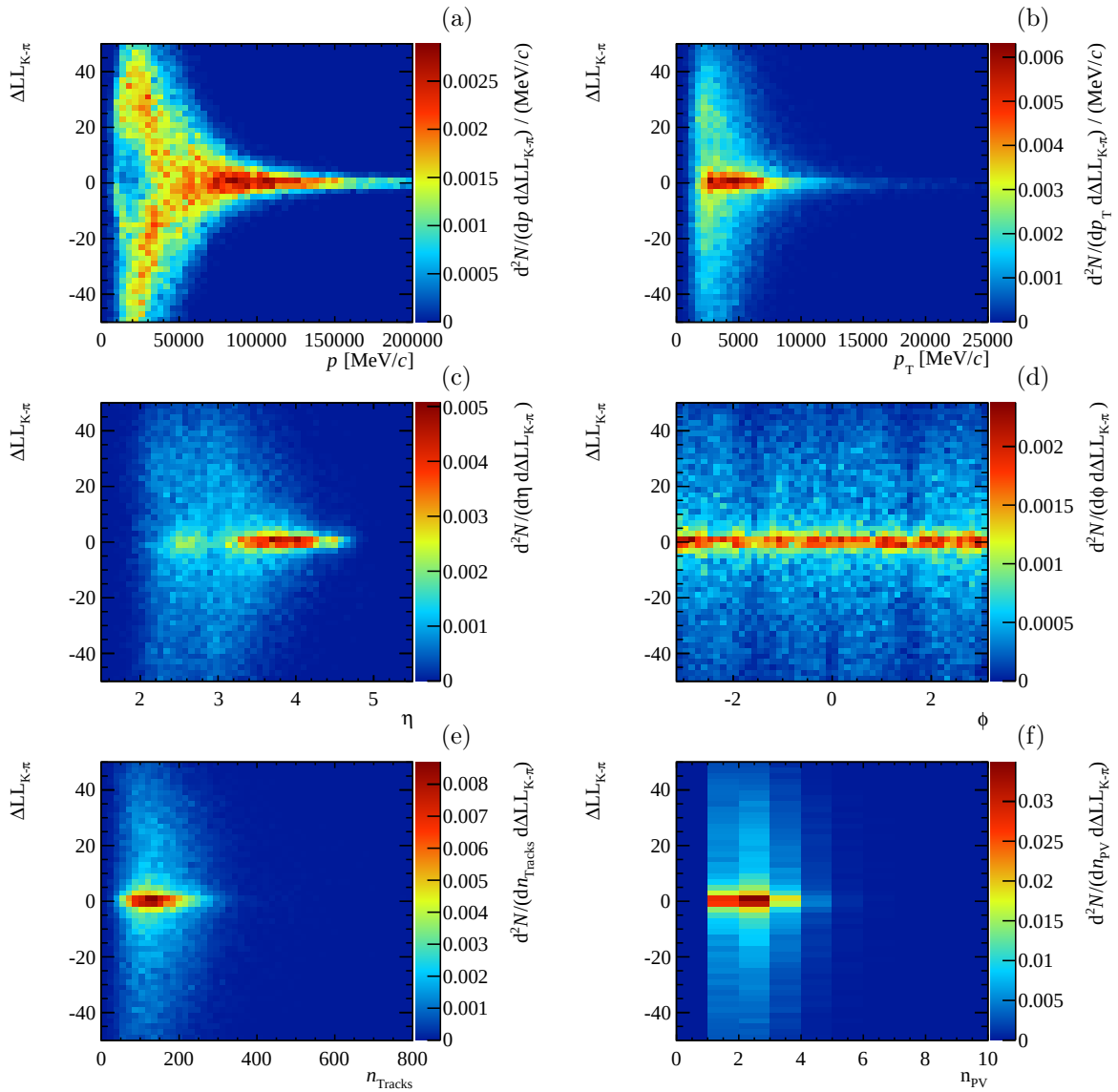


Figure 7.7: Dependence of $\Delta LL_{K-\pi}$ on (a) the track momentum, p; (b) the transverse momentum, p_T, of the track; (c) the track pseudorapidity, η; (d) the azimuthal angle, ϕ, of the track; (e) the number of tracks in the event, and (f) the number of primary vertices in the event. The distributions are taken from a sample of $B^0_{(s)} \to h^+h'^-$ candidates in data with BDTS > 0.5 and $5200 < m_{h^+h'^-} < 5400$ to reduce the background.

p: [5,9.3,17.675,15.6,20,23,26,29.65,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100,125, 150,200] GeV/c [6]
η: [1.5,2.5,3.0,3.5,4.0,5.0]
n_{Tracks}: [0,50,100,200,400,800]
and for various cuts on $\Delta\text{LL}_{K-\pi}$ for kaons as well as for pions.

7.3.2 Evaluation of the PID efficiency

The output of the efficiency calculation is checked by comparing it to an estimate using $K^{\pm}$ candidates from data samples of $B^{\pm} \to J/\psi K^{\pm}$ candidates, for the performance on single tracks, and from data samples of $B_s^0 \to J/\psi \phi$ candidates, for checking the appropriate behaviour in the presence of several tracks with $\Delta\text{LL}_{K-\pi}$ requirements. The correlation between the $\Delta\text{LL}_{K-\pi}$ values of different final state particles is shown in Fig. 7.8.

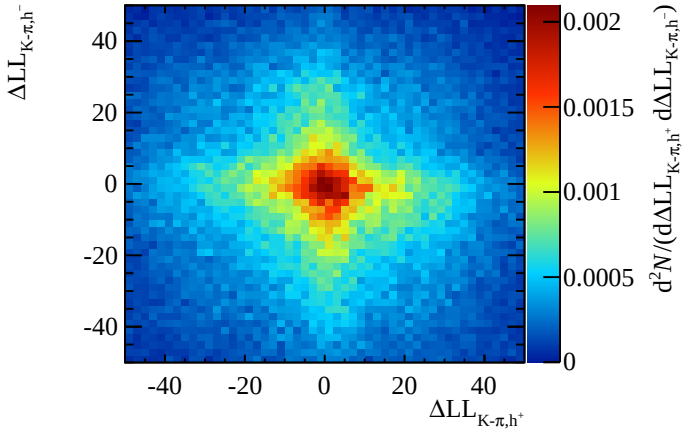


Figure 7.8: Distribution of the $\Delta\text{LL}_{K-\pi}$ values of the two hadrons in a sample of $B^0_{(s)} \to h^+h'^-$ candidates in data with BDTS > 0.5 and $5200 < m_{h^+h'^-} < 5400$ to reduce the background.

The comparison is done separately for the different stripping versions, S20 and S20r1 (2012 and 2011 data), as well as for the two magnet polarities.

Evaluation on $B^{\pm} \to J/\psi K^{\pm}$

The three-dimensional (p, η, number of tracks) efficiency maps obtained including the binomial uncertainties on the efficiency values for cuts on the $\Delta\text{LL}_{K-\pi}$ value of a kaon from the PIDCalibTools are folded with the normalised distributions of p and η of the kaon and the number of reconstructed tracks in the event of the $B^{\pm} \to J/\psi K^{\pm}$ data sample. Cuts of BDTS > 0.25 and $|m_{J/\psi} - m_{\mu^+\mu^-}| < 60$ MeV/c^2 have been applied to reduce the background in this $B^{\pm} \to J/\psi K^{\pm}$ sample to a minimum. The result of this folding is an average efficiency for kaons $\langle\varepsilon_{\text{kaon}}\rangle$ for a given $\Delta\text{LL}_{K-\pi}$ cut.
As a reference, the signal yields with ($N_{\text{sig,cut}}$) and without ($N_{\text{sig,tot}}$) the $\Delta\text{LL}_{K-\pi}$ cut are determined from a binned maximum-likelihood fit of the invariant $B^{\pm}$ candidate mass distribution.

[6] The bin boundaries at 9.3, 17.675, 15.6 and 29.65 GeV/c^2 are chosen due to the thresholds of the different radiators (Aerogel, C_4F_{10}, CF_4) in RICH1 and RICH2.

The invariant mass distribution is described by the sum of two Crystal Ball functions, ${}_d\mathcal{CB}$, and an exponential function in the invariant mass interval between 5150 and 5450 MeV/c^2. An example of such a fit is shown in Fig. 7.9.

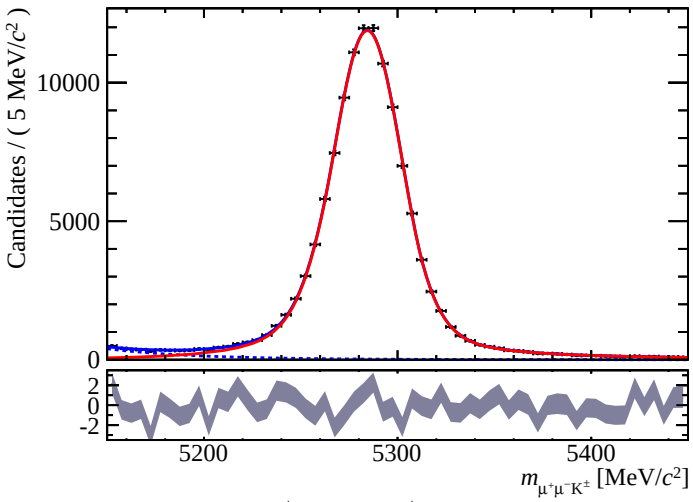


Figure 7.9: Invariant mass distribution of $B^\pm \to J/\psi K^\pm$ candidates without any PID cuts, described by the sum of two Crystal Ball functions for signal (red solid line) and an exponential function for background (blue dotted line). The full fit model is described by the solid blue line and the pull with respect to the data is shown by the grey band at the bottom.

The efficiency for this reference method is calculated as

$$\varepsilon_{\text{fit}} = \frac{N_{\text{sig,cut}}}{N_{\text{sig,tot}}}, \tag{7.3}$$

but as the signal yields from the fit include also candidates from the decay $B^\pm \to J/\psi \pi^\pm$ with a mis-identified pion ($\mathcal{B}(B^\pm \to J/\psi \pi^\pm)/\mathcal{B}(B^\pm \to J/\psi K^\pm) = 0.048$ [81]), the relation between ε_{fit} and $\langle \varepsilon_{\text{kaon}} \rangle$ is actually

$$\varepsilon_{\text{fit}} = \frac{\langle \varepsilon_{\text{kaon}} \rangle \mathcal{B}(B^\pm \to J/\psi K^\pm) + \langle \varepsilon_{\text{pion}} \rangle \mathcal{B}(B^\pm \to J/\psi \pi^\pm)}{\mathcal{B}(B^\pm \to J/\psi K^\pm) + \mathcal{B}(B^\pm \to J/\psi \pi^\pm)}, \tag{7.4}$$

and ε_{fit} can be corrected for the contribution of $B^\pm \to J/\psi \pi^\pm$ by calculating $\langle \varepsilon_{\text{pion}} \rangle$ in the same manner as $\langle \varepsilon_{\text{kaon}} \rangle$ based on the three-dimensional efficiency maps determined using the PIDCalibTools. This can be used to determine $\varepsilon_{\text{corr}}$

$$\varepsilon_{\text{corr}} = \varepsilon_{\text{fit}} + (\varepsilon_{\text{fit}} - \langle \varepsilon_{\text{pion}} \rangle) \frac{\mathcal{B}(B^\pm \to J/\psi \pi^\pm)}{\mathcal{B}(B^\pm \to J/\psi K^\pm)}, \tag{7.5}$$

which can be compared to the result from the PIDCalibTool $\langle \varepsilon_{\text{kaon}} \rangle$.
Figure 7.10 shows the results of the two methods as well as their difference. The difference $\Delta\varepsilon$ is not larger than 2 % over the considered range of $\Delta\text{LL}_{K-\pi}$ cuts. $\Delta\varepsilon$ as a function of the $\Delta\text{LL}_{K-\pi}$ cut is later used to evaluate systematic uncertainties.

As a cross check, the $B^{\pm} \to J/\psi\,\pi^{\pm}$ component is separated in the mass fit by using a Crystal Ball function to describe the $B^{\pm} \to J/\psi\,K^{\pm}$ component and a Gaussian function for $B^{\pm} \to J/\psi\,\pi^{\pm}$ as well as the exponential function for the combinatoiral background (*cf.* Fig. 7.11). The resulting efficiency values are in good agreement with the first method.

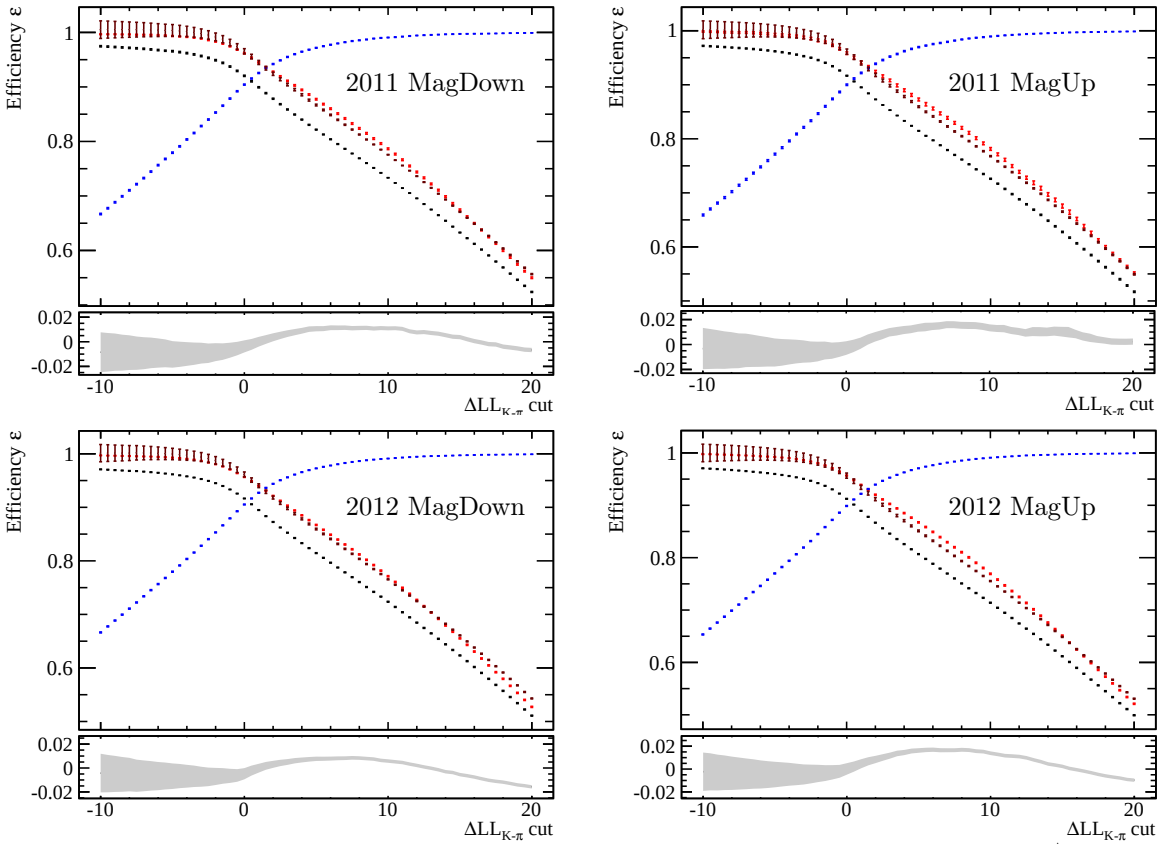


Figure 7.10: PID efficiency as a function of the $\Delta\mathrm{LL}_{K-\pi}$ cut on the kaon candidate in $B^{\pm} \to J/\psi\,K^{\pm}$ for 2011 (top) and 2012 (bottom) data as well as for the two magnet polarities. The bright red points show $\langle\varepsilon_{\mathrm{kaon}}\rangle$, the blue ones $1 - \langle\varepsilon_{\mathrm{pion}}\rangle$, the black ones $\varepsilon_{\mathrm{fit}}$ and the dark red ones $\varepsilon_{\mathrm{corr}}$. The grey bands show the difference between $\varepsilon_{\mathrm{corr}}$ and $\langle\varepsilon_{\mathrm{kaon}}\rangle$ and its uncertainty.

Evaluation on $B_s^0 \to J/\psi\,\phi$

The decay $B_s^0 \to J/\psi\,\phi$ is used to study the performance of the PIDCalibTool in the case of $\Delta\mathrm{LL}_{K-\pi}$ cuts on two tracks. Again a fit of the invariant B_s^0 candidate mass distribution with and without $\Delta\mathrm{LL}_{K-\pi}$ cuts on the kaon candidates is used to determine the average $\Delta\mathrm{LL}_{K-\pi}$ efficiency for kaons $\langle\varepsilon_{\mathrm{kaon}}\rangle$ as a function of the $\Delta\mathrm{LL}_{K-\pi}$ cut.
In this decay there are also background components, which must be taken into account. The major background component is the decay $B_s^0 \to J/\psi\,h^+h'^-$, *i.e.* the S-wave component in the dihadron system. As the branching fractions for $B_s^0 \to J/\psi\,h^+h'^-$, especially as a function of the dihadron mass, are not known, the different components are disentangled by a three-dimensional binned maximum-likelihood fit to the invariant B_s^0, J/ψ and ϕ candidate mass in the mass intervals $m_{\mu^+\mu^-K^+K^-} \in [5310, 5430]\,\mathrm{MeV}/c^2 \times m_{\mu^+\mu^-} \in [3050, 3150]\,\mathrm{MeV}/c^2 \times m_{K^+K^-} \in$

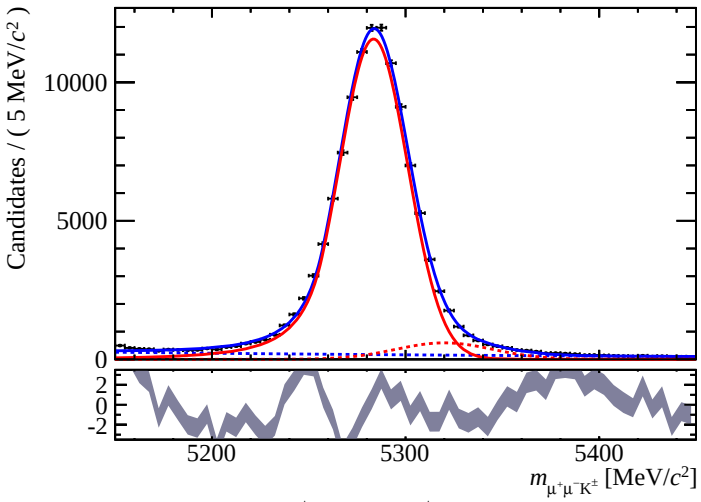


Figure 7.11: Invariant mass distribution of $B^\pm \to J/\psi\, K^\pm$ candidates without any PID cuts, described by the combination of a Crystal Ball functions for $B^\pm \to J/\psi\, K^\pm$ (red solid line), a Gaussian for $B^\pm \to J/\psi\, \pi^\pm$ (red dotted line) and an exponential function for background (blue dotted line). The full fit model is described by the solid blue line and the pull with respect to the data is shown by the grey band at the bottom.

$[1014.5, 1024.5]\,\mathrm{MeV}/c^2$. The idea is that the $B_s^0 \to J/\psi\,\phi$ component shows a resonance in all three mass variables while $B_s^0 \to J/\psi\, h^+h'^-$ only shows a resonance in the invariant B_s^0 and J/ψ candidate mass and the combinatorial background component has none at all. The background from the rare decay $B_s^0 \to \phi\mu^+\mu^-$ is neglected.

The signal and the $B_s^0 \to J/\psi\, h^+h'^-$ component in the B_s^0 candidate mass are each described by a ${}_d\mathcal{CB}$ functions. The two ${}_d\mathcal{CB}$ functions have the same central value and tail parameters, but different widths. The combinatorial background in this variable is described by an exponential function.
In the J/ψ candidate mass the signal and the $B_s^0 \to J/\psi\, h^+h'^-$ are described by two identical ${}_d\mathcal{CB}$ functions, the combinatorial background by an exponential function.
In the case of the ϕ candidate mass the signal is described by a Voigt function, *i.e.* a convolution of a Gaussian function describing the detector resolution and a Breit-Wigner function describing the significant decay width Γ of the ϕ. Γ is fixed to its nominal value of $4.26\,\mathrm{MeV}/c^2$ [81] in the fit. The components for $B_s^0 \to J/\psi\, h^+h'^-$ and the combinatorial background are each described by a square-root function

$$f(m_{K^+K^-}|m_0) = \sqrt{m_{K^+K^-} - m_0} \tag{7.6}$$

with the same parameter m_0.
The projections of the three-dimensional distributions for $B_s^0 \to J/\psi\,\phi$ candidates with BDTS $>$ 0.05 onto the three invariant masses $m_{\mu^+\mu^-K^+K^-}$, $m_{\mu^+\mu^-}$ and $m_{K^+K^-}$ are shown in Fig. 7.12.

The average efficiency of the PIDCalibTool $\langle\varepsilon_{\mathrm{kaon}}\rangle$ for a given $\Delta\mathrm{LL}_{K-\pi}$ cut is evaluated by calculating for each event the efficiency given by the efficiency tables according to the momenta

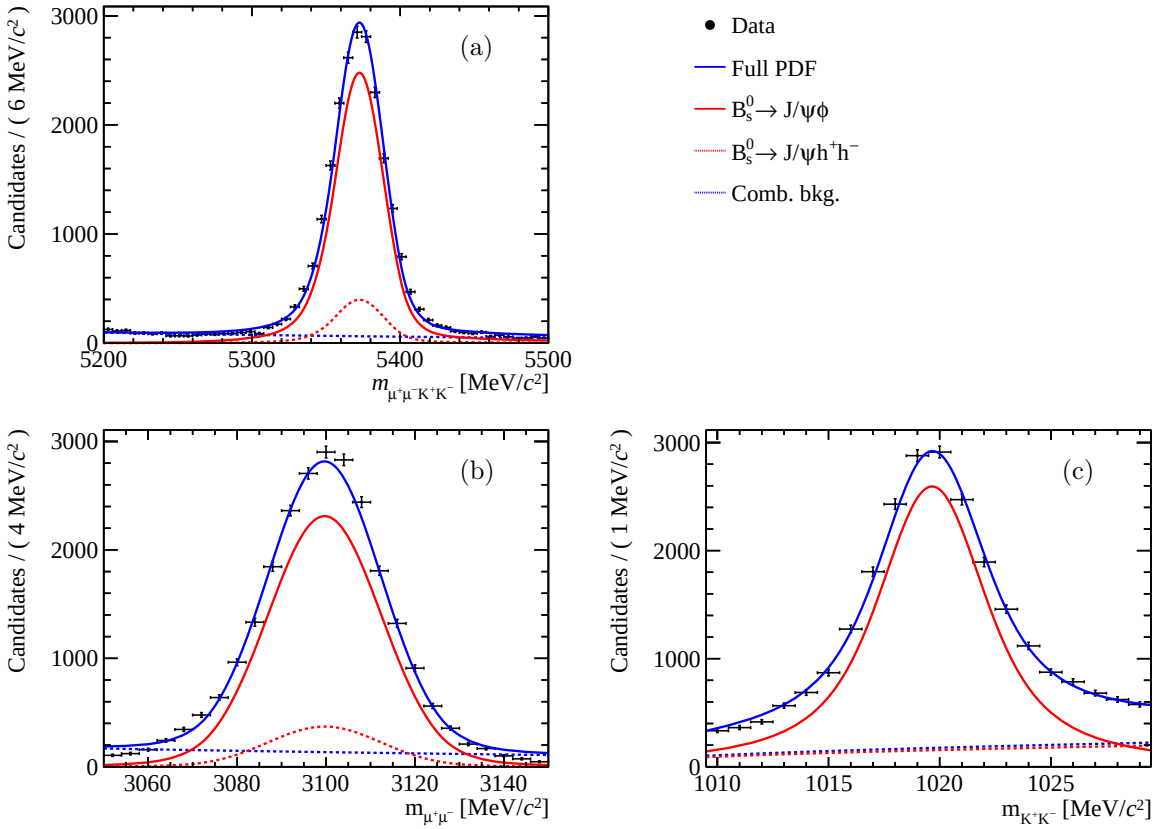


Figure 7.12: The invariant mass distributions for (a) $m_{\mu^+\mu^-K^+K^-}$, (b) $m_{\mu^+\mu^-}$ and (c) $m_{K^+K^-}$ used to extract the signal yield of $B_s^0 \to J/\psi\,\phi$. The signal component of the maximum-likelihood fit is shown by the red solid lines, the one of $B_s^0 \to J/\psi\, h^+h'^-$ by the red dashed lines and the one of the combinatorial background by the blue dashed lines.

and pseudorapidities of the two hadrons as well as the number of tracks in the event

$$\langle \varepsilon_{\text{kaon}} \rangle = \frac{\sum_{i=1}^{N_{\text{evt}}} w_i \cdot \varepsilon_{\text{kaon}}(p_{h_i^+}, \eta_{h_i^+}, \text{nTracks}) \cdot \varepsilon_{\text{kaon}}(p_{h_i'^-}, \eta_{h_i'^-}, \text{nTracks})}{\sum_{i=1}^{N_{\text{evt}}} w_i}, \tag{7.7}$$

where w_i is the signal weight extracted by the ${}_s\mathcal{P}lot$ method based on the fit of the invariant mass distributions described above to disentangle signal and background.
Figure 7.13 shows $\langle \varepsilon_{\text{kaon}} \rangle$, the efficiency estimated from the fit and the difference between the two. The result based on the output of the PIDCalibTool is within the uncertainties in perfect agreement with the estimation based on the fit of the invariant mass distributions.

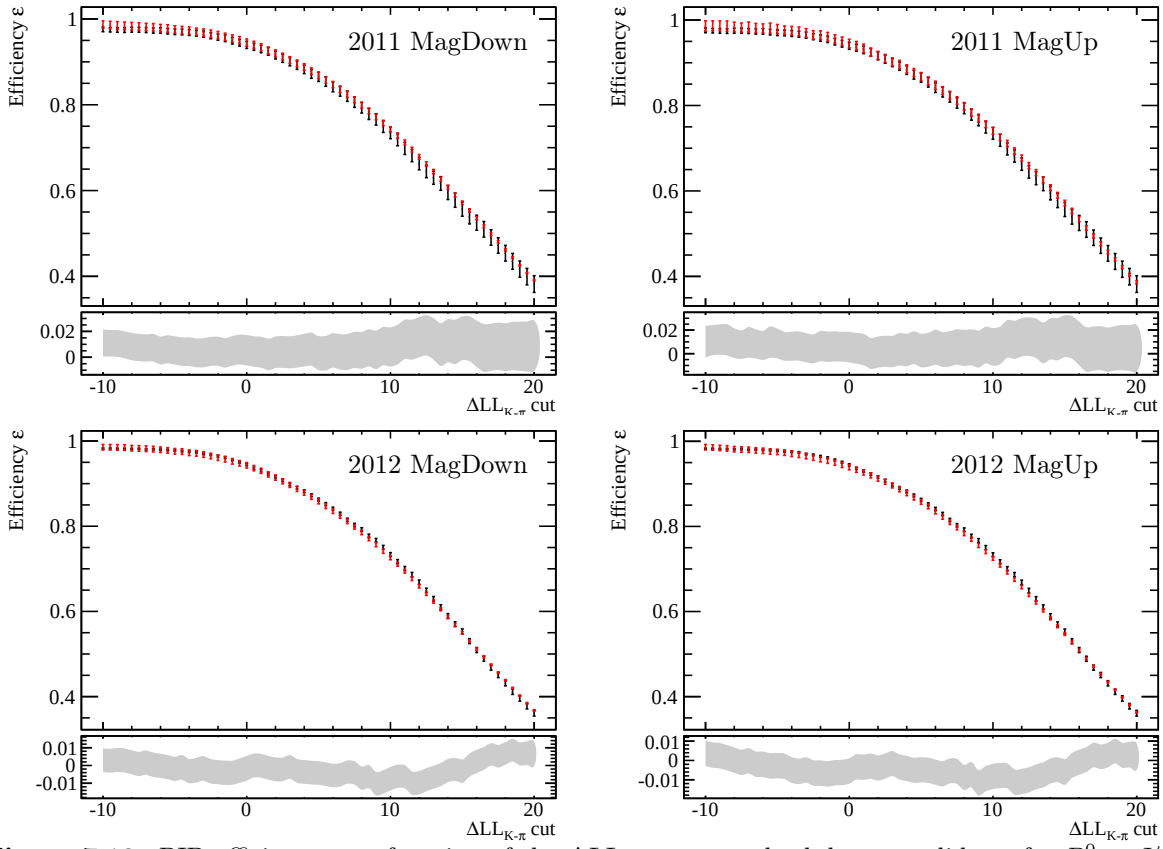


Figure 7.13: PID efficiency as a function of the $\Delta\text{LL}_{K-\pi}$ cut on both kaon candidates for $B_s^0 \to J/\psi\,\phi$ in 2011 (top) and 2012 (bottom) data as well as for both magnet polarities. The red points show the estimation by the PIDCalibTool while the black points show the estimation from a fit of the invariant mass distributions of the B_s^0, J/ψ and ϕ candidates. The grey bands at the bottom show the difference between the two methods and its uncertainty.

7.4 Calibration

The goal of the calibration is to determine the BDT distribution in its range between 0 and 1 for signal candidates. Thereto, the range of the BDT values is split into eight bins [0.00,0.25,0.40,0.50,

0.60,0.70,0.80,0.90,1.00]. In each of these eight bins the fraction of signal is determined by fitting the invariant mass distribution of the $B^0_{(s)} \to h^+h'^-$ candidates. By cuts on the $\Delta\mathrm{LL}_{K-\pi}$ variable of the two final state hadrons, the most probable mass hypothesis ($m_{\pi^+\pi^-}$, $m_{K^+\pi^-}$, $m_{\pi^+K^-}$ or $m_{K^+K^-}$) can be assigned to each candidate. Henceforth, $m_{h^+h'^-}$ refers to the invariant mass based on the assigned types of final state particles.
The variable $\kappa \geq 0$ is introduced to define the cuts on the $\Delta\mathrm{LL}_{K-\pi}$ variable. For a given value of κ a hadron is defined as a kaon if $\Delta\mathrm{LL}_{K-\pi} > \kappa$ and as a pion if $\Delta\mathrm{LL}_{K-\pi} < -\kappa$. If at least one hadron in the candidate has a $\Delta\mathrm{LL}_{K-\pi}$ value larger than $-\kappa$ but smaller than κ, the candidate is discarded.

7.4.1 Fit of the Invariant Dihadron Mass Distribution

The $B^0_{(s)} \to h^+h'^-$ candidates for the BDT calibration are selected as described in Tab. 6.1. In addition it is required that both final state hadrons lie within the acceptance of the Muon System to minimise possible biases between final state muons and hadrons. Further each final state hadron must have a p_T larger than 500 MeV/c, but smaller than 40 GeV/c. Finally, it is required that the event of the candidate is triggered independently of the signal by the L0 and HLT1 trigger levels[7]. The goal of this last requirement is to avoid biases implied by differences between hadronic (calibration channel) and muonic (signal channel) trigger lines. The HLT2 trigger level is not included in this requirement (*i.e.* no HLT2PhysTIS requirement) to have a large enough data sample.
The invariant dihadron mass distribution, $m_{h^+h'^-}$, is constructed by identifying each of the two final state particles as pion or kaon or discarding the candidate as described above. The accepted candidates are weighted by the inverse of the product of the PID efficiencies of the corresponding $\Delta\mathrm{LL}_{K-\pi}$ cuts

$$w = \left(\varepsilon_{\mathrm{kaon/pion}}(p_{h_i^+}, \eta_{h_i^+}, \mathrm{nTracks}) \cdot \varepsilon_{\mathrm{kaon/pion}}(p_{h_i'^-}, \eta_{h_i'^-}, \mathrm{nTracks})\right)^{-1}. \tag{7.8}$$

Candidates, in which one of the tracks has a $\Delta\mathrm{LL}_{K-\pi}$ efficiency smaller than 2 % or the relative uncertainty on the efficiency is larger than 50 %, are skipped to avoid the dominance of events with very large weights. This rejection affects even for the tightest $\Delta\mathrm{LL}_{K-\pi}$ cuts not more than 2 % of the candidates and the fraction of rejected candidates is stable across the different BDT bins.
The invariant dihadron mass distribution is studied for values of κ between 0.0 and 10.0 in steps of 0.5 to take into account a possible impact of the choice of the $\Delta\mathrm{LL}_{K-\pi}$ cut.

For each BDT bin, and for each value of κ, the signal yield is extracted from a binned maximum-likelihood fit to the invariant mass distribution. The fit model consists of two ${}_d\mathcal{CB}$ functions describing the B^0 and the B^0_s signals. All tail parameters ($\alpha_{l,r}$ and $n_{l,r}$) as well as the parameter f describing the relative fraction of the two ${}_d\mathcal{CB}$ functions are fixed to the values extracted from MC (*cf.* Tab. 7.1). The ratio of the widths of the functions describing the B^0_s and the B^0 signals is fixed to 1.0180 ± 0.0005 extracted by the interpolation method to estimate the mass resolution

[7] The actual requirement is that the candidate is classified as L0GlobalTIS and HLT1PhysTIS (TIS: Triggered independently on signal). In the case of a b hadron this can be for example achieved by the other b hadron from the $b\bar{b}$ pair firing the trigger. In contrast to GlobalTIS, PhysTIS only considers a subset of trigger lines (*e.g.* excluding dedicated lines for beam-gas collisions or luminosity measurements).

for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ described in Sec. 8.2.1.
Furthermore, the model includes an additional double Crystal Ball function where the two Crystal Ball functions have common central values, but different widths of the Gaussian part to take into account backgrounds from the decays $\Lambda_b^0 \to p\pi^-$ and $\Lambda_b^0 \to pK^-$ with the proton misidentified as a kaon or a pion. For this double Crystal Ball function, all parameters are fixed to the values extracted from MC (*cf.* Tab. 7.1) except for the central value, μ, which is allowed to vary in a $\pm 10\,\mathrm{MeV}/c^2$ window around the value determined in MC to take into account possible differences in the momentum scale between data and MC.
Further there is a phenomenological function describing background from partially reconstructed decays (*e.g.* $B^0 \to h^+h'^-\pi^0$)

$$\begin{aligned} f(m_{h^+h'^-}|m_0, c_p, \sigma_p) &= N \cdot \int_{-\infty}^{\infty} m' \Big(1 - \frac{m'^2}{m_0^2}\Big) \Theta(m_0 - m') e^{-c_p \cdot m'} \\ &\quad \cdot \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{(m-m')^2}{2\sigma_p^2}} \mathrm{d}m', \end{aligned} \tag{7.9}$$

where Θ is the Heaviside-Theta function and N is a normalisation factor [43, 105]. In the fit, all parameters of this function are free. The combinatorial background is described by an exponential function.

Figures 7.14, 7.15 and 7.16 show the invariant mass distributions in the different BDT bins and for different values of κ. The fraction $r_i(\kappa)$ of $B^0_{(s)} \to h^+h'^-$ events in the i-th BDT bin as a function of κ is calculated as

$$r_i(\kappa) = \frac{N_{B^0_{(s)},i}(\kappa)}{\sum_{j=1}^{8} N_{B^0_{(s)},j}(\kappa)}, \tag{7.10}$$

where $N_{B^0_{(s)},i}(\kappa)$ is the fitted signal yield in the i-th bin. Figure 7.17 shows $r_i(\kappa)$ for the eight BDT bins. For each BDT bin i, the weighted average $\overline{r}_i$ over the considered range of κ values is calculated to extract the BDT distribution for signal. As the data sample for a value κ' is a subset of the data sample for $\kappa < \kappa'$ there are strong correlations between $r_i(\kappa)$ and $r_i(\kappa')$, which are taken into account.
The correlation $\rho_{i;\kappa,\kappa'}$ between $r_i(\kappa)$ and $r_i(\kappa')$ is estimated from the number of unweighted candidates in the i-th BDT bin passing the $\Delta\mathrm{LL}_{K-\pi}$ cut, κ, $n_{i;\kappa}$ and those passing the cut, $\kappa' > \kappa$, $n_{i;\kappa'}$

$$\rho_{i;\kappa,\kappa'} = \sqrt{\frac{n_{i;\kappa'}}{n_{i;\kappa}}}. \tag{7.11}$$

This formula is derived by comparing the binomial uncertainty, σ_ε, on the ratio $\varepsilon = n_{i;\kappa'}/n_{i;\kappa}$ given by

$$\sigma_\varepsilon = \sqrt{\frac{\varepsilon(1-\varepsilon)}{n_{i;\kappa}}} \tag{7.12}$$

with the result of the Gaussian error propagation

$$\sigma_\varepsilon = \sqrt{\left(\frac{\partial\varepsilon}{\partial n_{i;\kappa'}}\right)^2 n_{i;\kappa'} + \left(\frac{\partial\varepsilon}{\partial n_{i;\kappa}}\right)^2 n_{i;\kappa} + 2\rho\left(\frac{\partial\varepsilon}{\partial n_{i;\kappa'}}\right)\left(\frac{\partial\varepsilon}{\partial n_{i;\kappa}}\right)\sqrt{n_{i;\kappa'} n_{i;\kappa}}}, \tag{7.13}$$

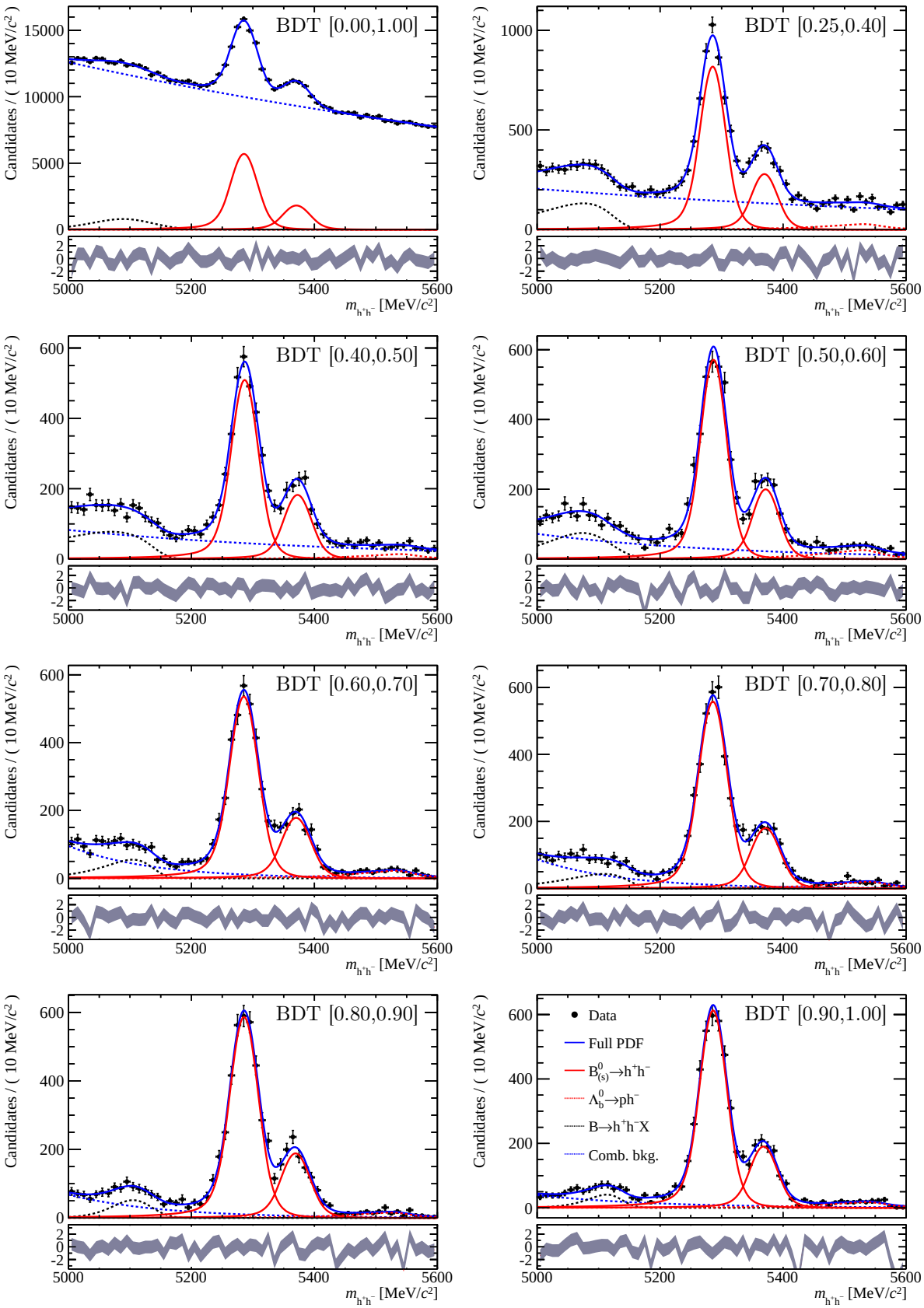


Figure 7.14: Invariant $m_{h^+h'^-}$ distribution for a $\Delta\mathrm{LL}_{K-\pi}$ cut value of $\kappa = 2$ in the highest seven BDT bins and for the full BDT range. The full fit model is shown by the blue solid line, the signal components by the red solid lines, the component from $\Lambda_b^0 \to ph^-$ by the red dashed line, the component from partially reconstructed background by the black dashed line and the combinatorial background by the blue dashed line.

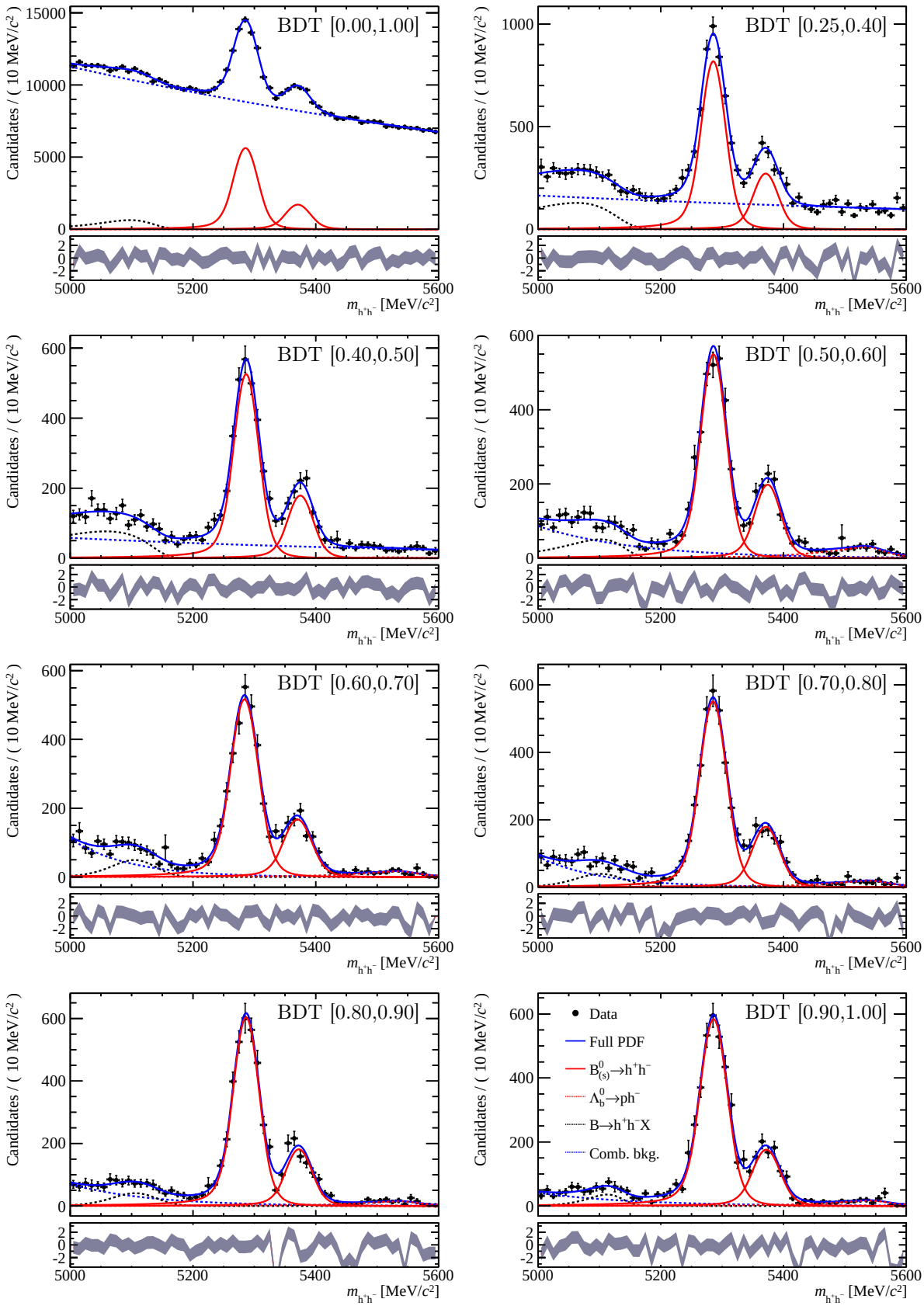


Figure 7.15: Invariant $m_{h^+h'^-}$ distribution for a $\Delta\mathrm{LL}_{K-\pi}$ cut value of $\kappa = 5$ in the highest seven BDT bins and for the full BDT range. The full fit model is shown by the blue solid line, the signal components by the red solid lines, the component from $\Lambda_b^0 \to ph^-$ by the red dashed line, the component from partially reconstructed background by the black dashed line and the combinatorial background by the blue dashed line.

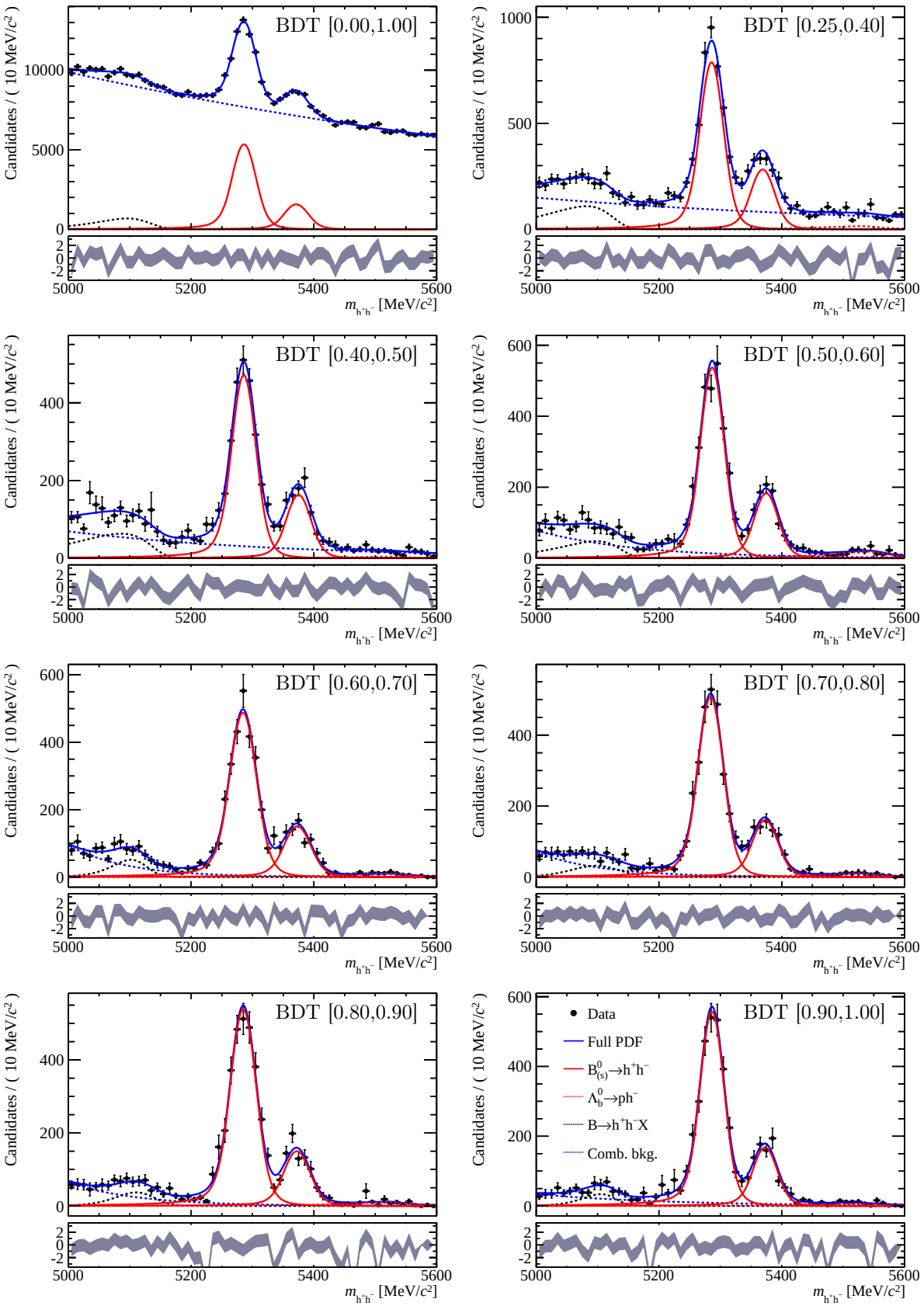


Figure 7.16: Invariant $m_{h^+h'^-}$ distribution for a $\Delta\mathrm{LL}_{K-\pi}$ cut value of $\kappa = 8$ in the highest seven BDT bins and for the full BDT range. The full fit model is shown by the blue solid line, the signal components by the red solid lines, the component from $\Lambda_b^0 \to ph^-$ by the red dashed line, the component from partially reconstructed background by the black dashed line and the combinatorial background by the blue dashed line.

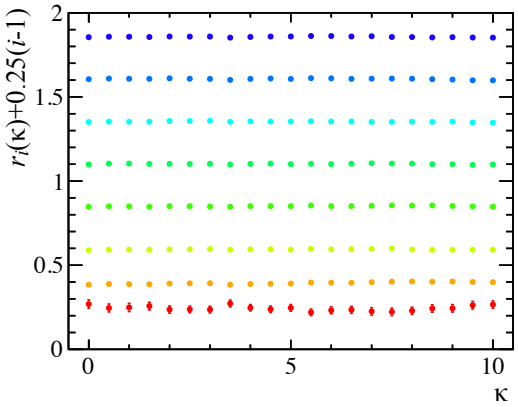


Figure 7.17: The fraction, r_i, of signal candidates for each BDT bin i as a function of the κ represented by the variable κ. For better visibility, the data points have a vertical offset by $0.25 \cdot (i-1)$ $(i = 1, \ldots, 8)$.

using $\sigma_{n_{i;\kappa}} = \sqrt{n_{i;\kappa}}$ and $\sigma_{n_{i;\kappa'}} = \sqrt{n_{i;\kappa'}}$.
The resulting fractions of the signal yield in each BDT bin together with their statistical and systematic uncertainties are listed in Tab. 7.2. The systematic uncertainties are discussed in the following subsection.

Table 7.2: Estimated fractions of $B^0_{(s)} \to h^+h'^-$ signal yields in the eight BDT bins (value) with their statistical (stat.) and systematic (syst.) uncertainties

BDT bin	**value stat. syst.**
$[0.00, 0.25]$	$0.2437 \pm 0.0095 \pm 0.0105$
$[0.25, 0.40]$	$0.1425 \pm 0.0038 \pm 0.0115$
$[0.40, 0.50]$	$0.0935 \pm 0.0028 \pm 0.0050$
$[0.50, 0.60]$	$0.1006 \pm 0.0034 \pm 0.0040$
$[0.60, 0.70]$	$0.1018 \pm 0.0033 \pm 0.0039$
$[0.70, 0.80]$	$0.1033 \pm 0.0031 \pm 0.0054$
$[0.80, 0.90]$	$0.1072 \pm 0.0029 \pm 0.0046$
$[0.90, 1.00]$	$0.1074 \pm 0.0027 \pm 0.0037$

7.4.2 Systematic Uncertainties

Several sources of systematic uncertainties in the determination of the BDT distribution for the signal have been studied. The details are discussed below and the estimated uncertainties from the different sources are listed in Tab. 7.3.

1. **Effect of the $\Delta LL_{K-\pi}$ cut value** The BDT distribution for signal should not depend on the value of κ. To check for changes in the fraction of the signal yield in the different BDT bins, a χ^2-fit of a linear function is performed to the signal fractions as a function of κ for each BDT bin taking into account the correlations among the signal fractions for different

κ values. These fits are shown in Fig. 7.18. The fraction of the signal yield for $\kappa = 0$ is extrapolated $r_{\mathrm{i;extra}}$ using the fitted linear function. If this value is significantly different from the weighted average $\overline{r}_i$ the absolute difference between the two values is taken as systematic uncertainty. Otherwise, no systematic uncertainty is assigned. Significant means that $|r_{i;\mathrm{extra}} - \overline{r}_i| > \sqrt{\sigma^2_{r_{i;\mathrm{extra}}} + \sigma^2_{\overline{r}_i}}$ where $\sigma_{r_{i;\mathrm{extra}}}$ is the uncertainty on the extrapolated value extracted from the fit and $\sigma_{\overline{r}_i}$ is the uncertainty on the weighted average $\overline{r}_i$.

2. **Binning scheme in the $\Delta\mathrm{LL}_{K-\pi}$ efficiency determination** Possible effects of the binning scheme in which the $\Delta\mathrm{LL}_{K-\pi}$ efficiency is determined are estimated by repeating the determination of the $\overline{r}_i$ using an alternative binning scheme for the $\Delta\mathrm{LL}_{K-\pi}$ efficiency tables. The alternative binning scheme consists of 46 bins in the track momentum, 7 bins in the track pseudorapidity and 7 bins in the number of tracks:

   p: [5,9.3,12,15.6,17.675,20,21.5,23,24.5,26,27.5,29.65,30,32.5,35,37.5,40,42.5,45,47.5, 50,52.5,55,57.5,60,62.5,65,67.5,70,72.5,75,77.5,80,82.5,85,87.5,90,92.5,95,97.5,100, 110,125,135,150,170,200] GeV/c
   η: [1.5,2.0,2.5,3.0,3.5,4.0,4.5,5.0]
   n_{Tracks}: [0,50,75,100,150,200,400,800]

   The difference between the values of $\overline{r}_i$ obtained using this binning scheme and the nominal scheme is taken as systematic uncertainty.

3. **Uncertainty on the $\Delta\mathrm{LL}_{K-\pi}$ efficiency** The uncertainty induced by the $\Delta\mathrm{LL}_{K-\pi}$ efficiency is estimated by changing the event-by-event weighting to

$$w = \left([\varepsilon_{\mathrm{kaon/pion}}(p_{h_i^+}, \eta_{h_i^+}, \mathrm{n_{Tracks}}) + \Delta\varepsilon(\kappa)] \cdot [\varepsilon_{\mathrm{kaon/pion}}(p_{h_i'^-}, \eta_{h_i'^-}, \mathrm{n_{Tracks}}) + \Delta\varepsilon(\kappa)] \right)^{-1}, \tag{7.14}$$

   where $\Delta\varepsilon(\kappa)$ is the uncertainty on the $\Delta\mathrm{LL}_{K-\pi}$ efficiency estimated with $B^\pm \to J/\psi K^\pm$ candidates (*cf.* Sec. 7.3.2). The difference between the values for $\overline{r}_i$ obtained using this modified weighting and the nominal weighting is assigned as systematic uncertainty.

4. **Skipped candidates** As mentioned above, candidates where one track has a very large relative uncertainty on the $\Delta\mathrm{LL}_{K-\pi}$ efficiency or a very small efficiency are skipped. The largest fraction of skipped candidates as a function of κ is not larger than 2 %. Further, there is also no systematic difference between the BDT bins in the fraction of skipped candidates. As systematic uncertainty the fraction of signal yields is recalculated assuming the case that all skipped candidates are signal. The differences with respect to the nominal values of $\overline{r}_i$ are taken as systematic uncertainty.

5. **Fit model of the $m_{h^+h'^-}$ distribution** The impact of the choice of the fit model describing the $m_{h^+h'^-}$ distribution is estimated by repeating the binned maximum-likelihood fits of the $m_{h^+h'^-}$ distributions with a modified fit model. The ${}_d\mathcal{CB}$ functions to describe the B^0 and B_s^0 signals were each replaced by the sum of two Gaussians with different widths, but common mean. The combinatorial background is not described by a single exponential function, but by the sum of two exponential functions. Figure 7.19 shows the fitted $m_{h^+h'^-}$ distributions.
   The fractions of signal in the BDT bins is calculated based on the signal yields extracted

from this fit model and the differences to the nominal fractions shown in Tab. 7.2 are taken as systematic uncertainty.

The total systematic uncertainties on the fractions of signal candidates are calculated as the square root of the squared sum of the individual uncertainties from the five sources listed above.

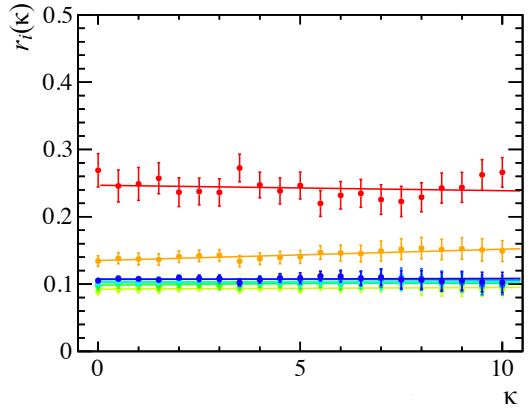


Figure 7.18: The linear regression of the fraction of signal candidates for each BDT bin as a function of κ. The different colours represent the different BDT bins (1: red, 2: orange, 3: yellow-green, 4: green, 5: darkgreen, 6: cyan, 7: blue 8: darkblue).

Table 7.3: Systematic uncertainties on $\overline{r}_i$ in the different BDT bins

BDT bin	$\Delta LL_{K-\pi}$ **consistency**	$\Delta LL_{K-\pi}$ **binning scheme**	$\Delta LL_{K-\pi}$ **uncertainty**	**Skipped candidates**	$m_{h^+h'^-}$ **model**	**Total**
$[0.00, 0.25]$	0.0053	0.0035	0.0040	0.0053	0.0050	0.0105
$[0.25, 0.40]$	0.0093	0.0024	0.0041	0.0023	0.0039	0.0115
$[0.40, 0.50]$	0.0000	0.0022	0.0032	0.0012	0.0029	0.0050
$[0.50, 0.60]$	0.0000	0.0016	0.0017	0.0010	0.0031	0.0040
$[0.60, 0.70]$	0.0000	0.0015	0.0024	0.0015	0.0023	0.0039
$[0.70, 0.80]$	0.0028	0.0025	0.0027	0.0014	0.0024	0.0054
$[0.80, 0.90]$	0.0032	0.0014	0.0013	0.0021	0.0017	0.0046
$[0.90, 1.00]$	0.0000	0.0009	0.0022	0.0018	0.0021	0.0037

Figure 7.20 shows the BDT distribution for signal in the eight BDT bins together with the statistical and systematic uncertainties.

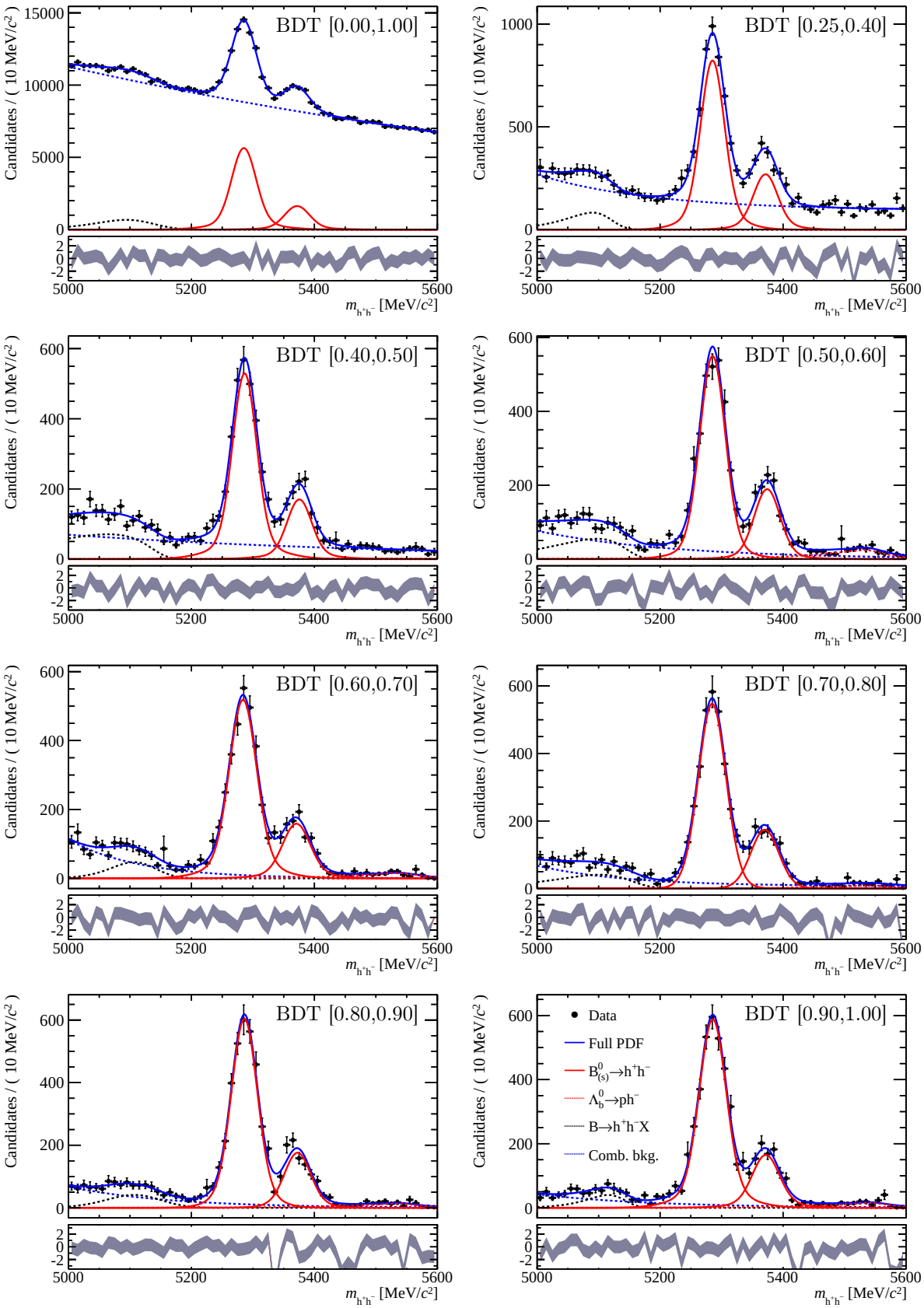


Figure 7.19: Invariant $m_{h^+h'^-}$ distribution for a $\Delta\mathrm{LL}_{K-\pi}$ cut value of $\kappa = 5$ in the highest seven BDT bins and for the full BDT range fitted with an alternative mass model described in the text. The full fit model is shown by the blue solid line, the signal components by the red solid lines, the component from $\Lambda_b^0 \to ph^-$ by the red dashed line, the component from partially reconstructed background by the black dashed line and the combinatorial background by the blue dashed line.

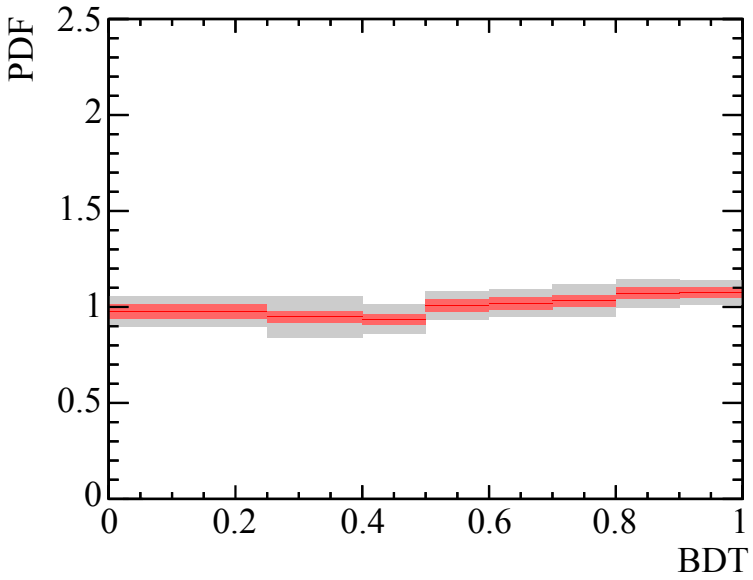


Figure 7.20: the signal PDF of the BDT distribution obtained from the calibration. It corresponds to $\overline{r}_i$ normalised to the width of the BDT bin. The red line shows the central value while the red and grey bands show the statistical and systematic uncertainties, respectively.

8

The Calibration of the Invariant Dimuon Mass for Signal

The invariant mass distribution for signal candidates in the analysis of $B^0_{(s)} \to \mu^+\mu^-$ is described by a single Crystal Ball distribution $\mathcal{CB}$ for B^0 and B^0_s each as described in Eq. (7.2). This distribution is characterised by the four parameters μ, σ, α, and n. The limited statistics in the signal channel requires that the parameters are estimated from calibration channels in data or from simulation.
Section 8.1 describes the estimation of the central value μ while Sec. 8.2 discusses the estimation of the width σ of the Gaussian kernel. The estimation of the tail parameters α and n is described in Sec 8.3. There also an alternative method is introduced to describe the additional parameters if double Crystal Ball distributions ${}_d\mathcal{CB}$ (*cf.* Eq. (7.1)) are used to describe the signal distributions. The results presented in this chapter are an update of those described in Ref. [137].

8.1 Central Value

The parameter μ of the Crystal Ball function describes the mean of the Gaussian kernel. Systematic biases in the momentum scale of the LHCb detector cause systematic shifts in the reconstructed invariant mass. This requires that μ is estimated from data rather than taking the nominal values of the masses of B^0 and B^0_s mesons. As for the BDT calibration described in the previous chapter, the decays $B^0_{(s)} \to h^+h'^-$ are used as a proxy for the decays $B^0_{(s)} \to \mu^+\mu^-$ to estimate μ from data.

8.1.1 Selection

The $B^0_{(s)} \to h^+h'^-$ candidates are selected in the same manner as described in the previous chapter (*cf.* Tab. 6.1). The only changes are a tighter cut on the BDTS variable (BDTS > 0.1) and the following cuts on the ΔLL classifiers: a charged hadron is identified as a kaon if $\Delta\mathrm{LL}_{K-\pi} > 10$ as well as $\Delta\mathrm{LL}_{K-p} > 2$ and as a pion if $\Delta\mathrm{LL}_{\pi-K} > 10$ as well as $\Delta\mathrm{LL}_{\pi-p} > 2$. There is no

requirement that the events are triggered independently of the $B^0_{(s)} \to h^+h'^-$ candidate.
The four final state hypotheses $\pi^+\pi^-$, $K^+\pi^-$, π^+K^- and K^+K^- are treated separately in the estimation of the central value.
In contrast to the invariant mass resolution, no impact on the central value of the resonances is expected by cuts on the ΔLL variables to first order. Nevertheless, a possible bias is – as described later – taken into account by one of the assigned systematic uncertainties.

8.1.2 Fit Model and Results

Results for the Central Value

Figures 8.1 and 8.2 show the invariant dihadron mass distributions of the candidates in the four final states $\pi^+\pi^-$, $K^+\pi^-$, π^+K^- and K^+K^- in 2011 and 2012 data, respectively.
The mass distributions are fit using a binned maximum-likelihood method to extract the central values for the B^0 and B^0_s signals. The fit model is similar to the one used in the calibration of the multivariate classifier described in the previous chapter. It consists of two ${}_d\mathcal{CB}$ functions – one for the B^0 and one for the B^0_s signal. The tail parameters as well as f in the two functions are fixed to the values determined from $B^0 \to K^\pm\pi^\mp$ MC (*cf.* Tab. 7.1). Further, the fit model includes the distribution described in Eq. (7.9) to describe physical background from partially reconstructed B decays and an exponential distribution describing combinatorial background.
A component for the decays $\Lambda^0_b \to ph^-$, where the proton is mis-identified, is not necessary due to the rejection of these decay modes due to the ΔLL cuts.
The extracted values of μ_{B^0} and $\mu_{B^0_s}$ are listed in Tab. 8.1. Also the ratio f_{B^0} of the fitted B^0 yield to the total B^0 + B^0_s yield is listed for the $K^+\pi^-$ and π^+K^- samples.
First of all, there is good agreement between 2011 and 2012 data. In the case of the B^0 meson, the obtained central values are about 5.5 MeV/c^2 above the nominal mass value [81], in the case of the B^0_s meson, they are about 5.2 MeV/c^2 above the nominal mass [81].

Table 8.1: The central values for two-body decays of B^0 (μ_{B^0}) and B^0_s ($\mu_{B^0_s}$) evaluated with the decays $B^0_{(s)} \to h^+h'^-$. Further the ratio of the B^0 to the total $B^0_{(s)}$ yield (f_{B^0}) is listed. The quoted uncertainties are statistical.

Decay channel	$\boldsymbol{\mu_{B^0}}$ **[MeV/c^2]**	$\boldsymbol{\mu_{B^0_s}}$ **[MeV/c^2]**	$\boldsymbol{f_{B^0}}$
2011 Data			
$B^0_{(s)} \to \pi^+\pi^-$	5284.99±0.52	–	–
$B^0_{(s)} \to \pi^+K^-$	5284.79±0.27	5370.3 ±5.4	0.951±0.006
$B^0_{(s)} \to K^+\pi^-$	5284.98±0.24	5373.6 ±6.5	0.923±0.005
$B^0_{(s)} \to K^+K^-$	–	5371.70±0.26	–
2012 Data			
$B^0_{(s)} \to \pi^+\pi^-$	5284.78±0.57	–	–
$B^0_{(s)} \to \pi^+K^-$	5284.96±0.19	5372.4 ±3.2	0.958±0.009
$B^0_{(s)} \to K^+\pi^-$	5284.85±0.17	5377.7 ±5.1	0.922±0.007
$B^0_{(s)} \to K^+K^-$	–	5371.96±0.22	–

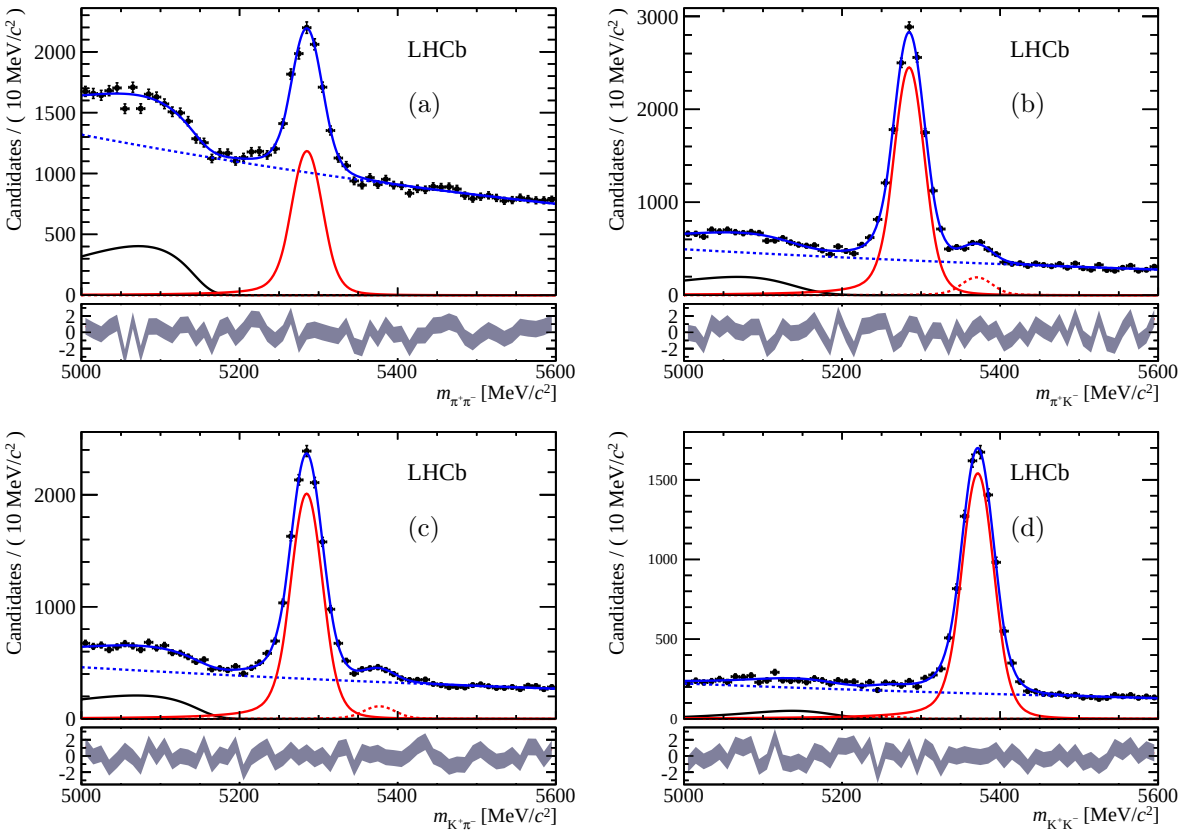


Figure 8.1: The invariant dihadron mass spectra for the decays $B^0_{(s)} \to h^+h'^-$ in 2011 data: (a) $\pi^+\pi^-$, (b) π^+K^-, (c) $K^+\pi^-$ and (d) K^+K^- candidates. The blue solid line shows the total fit model, the red lines show the B^0 and B^0_s signal (dominating signal: solid, sub-dominating signal: dotted), the black solid line shows the background from partially reconstructed B decays and the blue dashed line the combinatorial background.

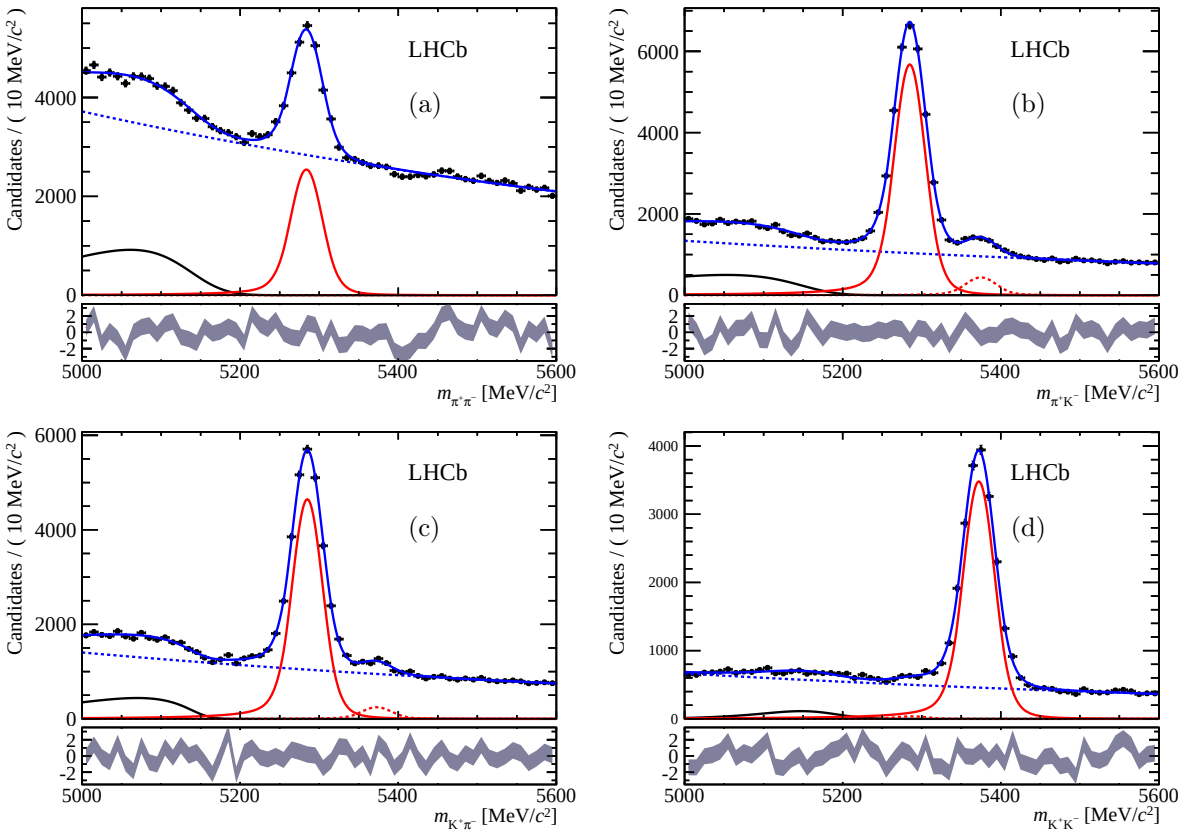


Figure 8.2: The invariant dihadron mass spectra for the decays $B^0_{(s)} \to h^+h'^-$ in 2012 data: (a) $\pi^+\pi^-$, (b) π^+K^-, (c) $K^+\pi^-$ and (d) K^+K^- candidates. The blue solid line shows the total fit model, the red lines show the B^0 and B^0_s signal (dominating signal: solid, sub-dominating signal: dotted), the black solid line shows the background from partially reconstructed B decays and the blue dashed line the combinatorial background.

The value of μ_{B^0} for the $B^0_{(s)} \to \mu^+\mu^-$ analysis is determined as the weighted average of the three results from the $\pi^+\pi^-$, π^+K^- and $K^+\pi^-$ final states while the value of $\mu_{B^0_s}$ from $B^0_{(s)} \to K^+K^-$ is used for $\mu_{B^0_s}$, as the other two results have large uncertainties.

Discussion of the Dihadron Spectra

The fitted yield of physical background is largest in the $\pi^+\pi^-$ sample and smallest in the K^+K^- sample. This reflects the relative abundance of pions and kaons in multi-body B decays.
The observed B^0 and B^0_s yields reflect the decay mechanisms in the different decay channels. The measured branching fractions from literature are listed in Tab. 8.2. The decays $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B^0_s \to \pi^+K^+$ and $B^0_s \to K^+K^-$ are allowed at tree level (*cf.* Fig. 8.3), but are suppressed by a factor $|V_{ub}|^2$ due to the involved $b \to u$ transition.

Table 8.2: World's average of the branching fractions for the decays $B^0_{(s)} \to h^+h'^-$ [81].

Decay channel	$\mathcal{B}(\boldsymbol{B^0 \to h^+h'^-})$	$\mathcal{B}(\boldsymbol{B^0_s \to h^+h'^-})$
$\pi^+\pi^-$	$(5.12\pm0.19) \times 10^{-6}$	$(7.6\ \pm1.9)\ \times 10^{-7}$
π^+K^-	–	$(5.5\ \pm0.6)\ \times 10^{-6}$
$K^+\pi^-$	$(1.96\pm0.05) \times 10^{-5}$	–
K^+K^-	$(1.3\ \pm0.5)\ \times 10^{-7}$	$(2.52\pm0.17) \times 10^{-5}$

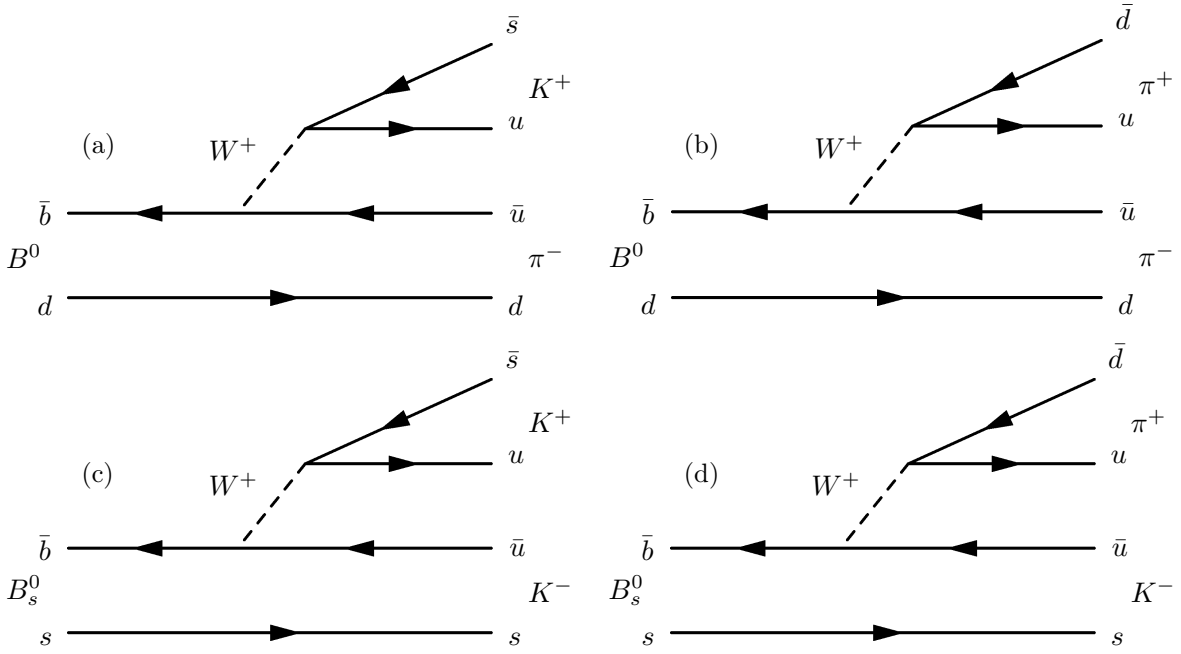


Figure 8.3: Tree level diagrams for the decays $B^0_{(s)} \to h^+h'^-$: (a) $B^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $B^0_s \to \pi^+K^+$ and (d) $B^0_s \to K^+K^-$ (N.B: All diagrams are suppressed by the CKM-matrix element V_{ub}).

In the case of the $\pi^+\pi^-$ final state, the $B^0 \to \pi^+\pi^-$ has a further contribution from internal

Penguin diagrams (*cf.* Figs 8.4 and 8.5). These diagrams have also a suppression by $|V_{ub}|^2$ and additionally in the Photon Penguin diagrams there is a color-suppression as the color of the quark-pair emitted from the photon has to match the one of the other quarks. On the other hand the decay $B_s^0 \to \pi^+\pi^-$ is only allowed via exchange diagrams. Exchange diagram amplitudes are typically suppressed by a factor 10 [147]. This leads to the absence of any significant B_s^0 signal in the $\pi^+\pi^-$ sample.

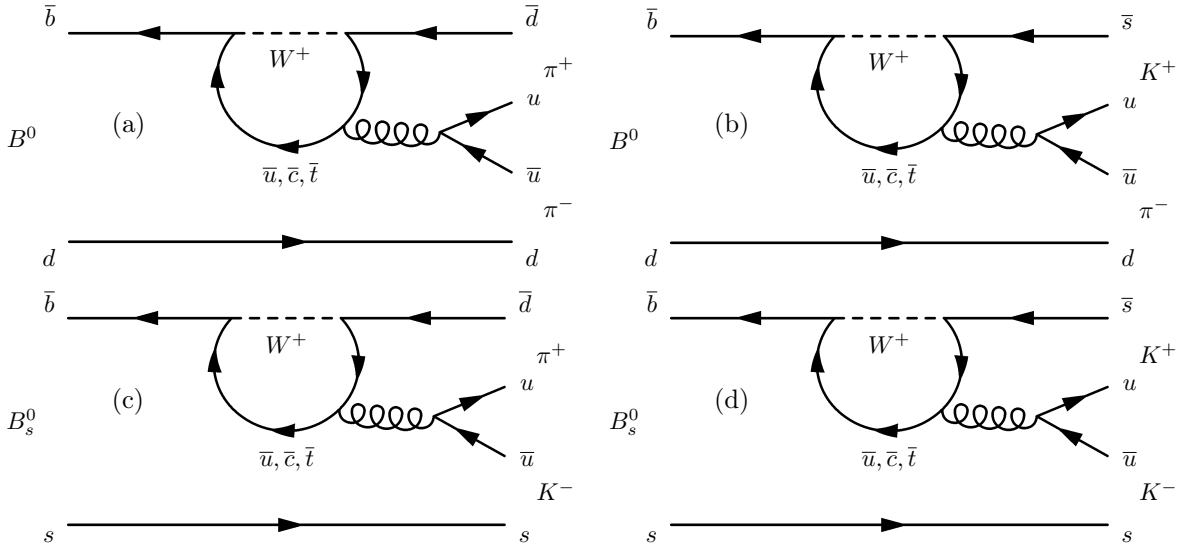


Figure 8.4: Gluon Penguin diagrams for the decays $B^0_{(s)} \to h^+h'^-$: (a) $B^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $B_s^0 \to \pi^+K^+$ and (d) $B_s^0 \to K^+K^-$.

For K^+K^-, the decay $B_s^0 \to K^+K^-$ also has contributions from internal Penguin diagrams. But in contrast to $B^0 \to \pi^+\pi^-$, these diagrams only suffer from a smaller Cabbibo-suppression by $|V_{ts}|^2$ explaining $\mathcal{B}(B_s^0 \to K^+K^-) > \mathcal{B}(B^0 \to \pi^+\pi^-)$. The absence of a B^0 signal in the K^+K^- sample is again explained by the exchange diagram for the decay $B^0 \to K^+K^-$ at leading order. Finally, the observed mass spectra for $K^+\pi^-$ and π^+K^- show a significantly smaller B_s^0 than B^0 signal. This is a result of the hadronisation ratio between B^0 and B_s^0 mesons $f_s/f_d = 0.259 \pm 0.015$ [26], but reflects also the larger branching fraction for $B^0 \to K^+\pi^-$ than for $B_s^0 \to \pi^+K^-$. The latter is a consequence that the first decay is – besides via the suppressed tree diagram – also allowed via an internal Penguin diagram suppressed only by $|V_{ts}|^2$ while the corresponding diagram for the second decay is also suppressed by $|V_{td}|^2$.

8.1.3 Systematic Uncertainties

Three sources of systematic uncertainties on μ_{B^0} and $\mu_{B_s^0}$ are considered: these are the cut on the BDTS variable as well as the two cuts on $\Delta\mathrm{LL}_{K-\pi}$ and $\Delta\mathrm{LL}_{K,\pi-p}$.

1. **BDTS** The central values are additionally determined using BDTS cuts at 0.20, 0.15 and 0.05. A linear regression to the central values extracted from these three samples and that

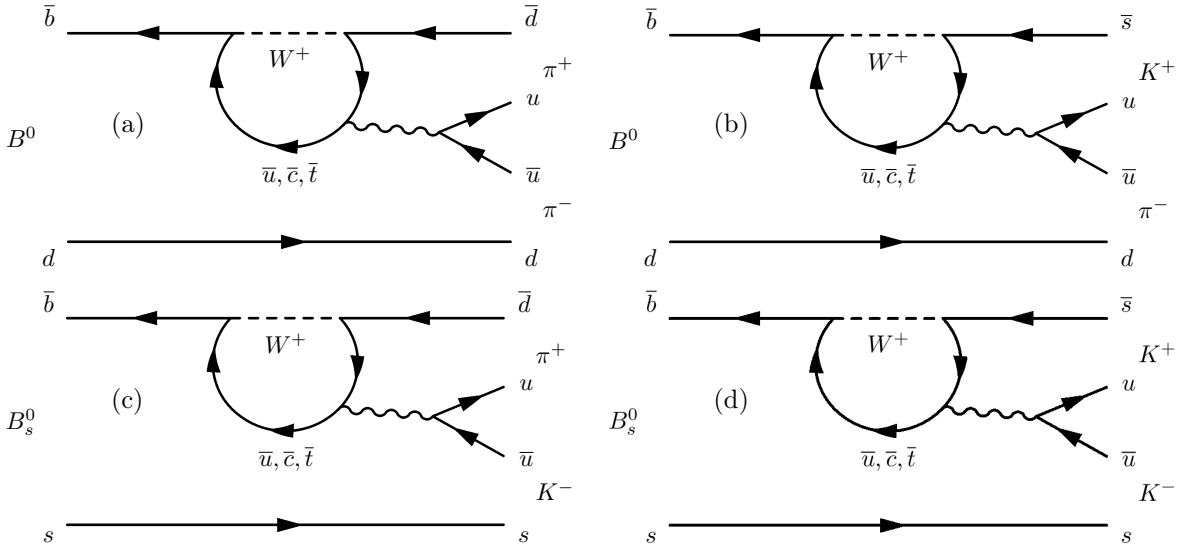


Figure 8.5: Photon Penguin diagrams for the decays $B^0_{(s)} \to h^+h'^-$: (a) $B^0 \to \pi^+\pi^-$, (b) $B^0 \to K^+\pi^-$, (c) $B^0_s \to \pi^+K^+$ and (d) $B^0_s \to K^+K^-$.

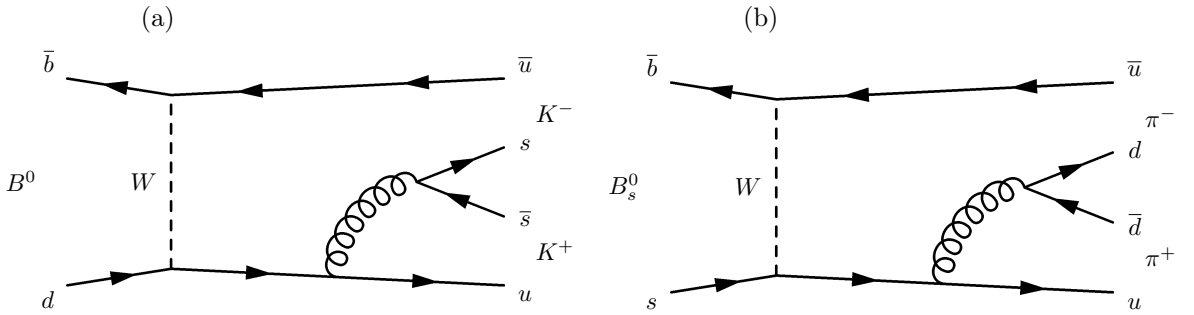


Figure 8.6: Exchange diagrams for the decays $B^0_{(s)} \to h^+h'^-$: (a) $B^0 \to K^+K^-$ and (b) $B^0_s \to \pi^+\pi^-$.

for BDTS > 0.10 is performed in each decay channel and the change of the linear function over the range of the BDTS cuts $[0.05, 0.20]$ is assigned as systematic uncertainty.

2. $\mathbf{\Delta LL_{K-\pi}}$ The determination of the central values is repeated with cuts on $\Delta\mathrm{LL}_{K-\pi}$ of 8, 9, 11 and 12 for kaons and -8, -9, -11 and -12 for pions, respectively. A linear regression using the central values extracted from these four samples and that for the nominal cuts is performed in each decay channel. The change of the linear function in the range of the $\Delta\mathrm{LL}_{K-\pi}$ cuts $[\pm 8, \pm 12]$ is assigned as systematic uncertainty.

3. $\mathbf{\Delta LL_{K,\pi-p}}$ The same procedure as for $\Delta\mathrm{LL}_{K-\pi}$ is employed with cut values 0, 1, 3 and 4 and considering the change of the fitted linear function in the range of the $\Delta\mathrm{LL}_{K,\pi-p}$ cuts $[0, 4]$ is assigned as systematic uncertainty.

The assigned systematic uncertainties are listed in Tab. 8.3. The total systematic uncertainty is calculated as the square-root of the sum of the squared contributions.

Table 8.3: Systematic uncertainties on the central values for two-body decays of B^0 (μ_{B^0}) and B_s^0 ($\mu_{B_s^0}$) evaluated with the decays $B^0_{(s)} \to h^+h'^-$.

Decay channel	**Systematic uncertainties [MeV/c^2]**			
	BDTS	$\mathbf{\Delta LL_{K-\pi}}$	$\mathbf{\Delta LL_{K,\pi-p}}$	**Total**
2011 Data				
$B^0 \to \pi^+\pi^-$	0.22	0.05	0.09	0.24
$B^0 \to \pi^+K^-$	0.14	0.04	0.06	0.16
$B^0 \to K^+\pi^-$	0.19	0.06	0.06	0.21
$B^0_{(s)} \to K^+K^-$	0.12	0.05	0.10	0.16
2012 Data				
$B^0 \to \pi^+\pi^-$	0.24	0.10	0.07	0.27
$B^0 \to \pi^+K^-$	0.19	0.08	0.08	0.22
$B^0 \to K^+\pi^-$	0.15	0.07	0.07	0.18
$B^0_s \to K^+K^-$	0.18	0.05	0.09	0.21

The central value for B_s^0 is that extracted from the $B_s^0 \to K^+K^-$ while for B^0 the weighted average over the three channels $B^0 \to \pi^+\pi^-$, $B^0 \to \pi^+K^-$ and $B^0 \to K^+\pi^-$ is calculated using the inverse of the squared statistical uncertainty as weighting factor. The total systematic uncertainty on μ_{B^0} is calculated as the weighted average of the individual systematic uncertainties also using the inverse of the squared statistical uncertainty as weighting factor.
The final results are

$$\begin{aligned} \mu_{B^0} &= 5284.91 \pm 0.17(\text{stat.}) \pm 0.19(\text{syst.})\,\text{MeV}/c^2 \\ \mu_{B_s^0} &= 5371.70 \pm 0.26(\text{stat.}) \pm 0.16(\text{syst.})\,\text{MeV}/c^2 \end{aligned}$$

for 2011 data and

$$\begin{aligned} \mu_{B^0} &= 5284.89 \pm 0.12(\text{stat.}) \pm 0.22(\text{syst.})\,\text{MeV}/c^2 \\ \mu_{B_s^0} &= 5371.96 \pm 0.22(\text{stat.}) \pm 0.22(\text{syst.})\,\text{MeV}/c^2 \end{aligned}$$

for 2012 data. As the extracted values are in good agreement between the 2011 and 2012 data, a common average is calculated, which is used in the description of the invariant dimuon mass distribution for the total data sample. Again, a weighted average is used based on the statistical uncertainty – the systematic uncertainties are correlated among the 2011 and 2012 data samples –, which gives

$$\begin{aligned} \mu_{B^0} &= 5284.90 \pm 0.10(\text{stat.}) \pm 0.21(\text{syst.})\,\text{MeV}/c^2 \\ \mu_{B_s^0} &= 5371.85 \pm 0.17(\text{stat.}) \pm 0.19(\text{syst.})\,\text{MeV}/c^2. \end{aligned}$$

8.2 Invariant Mass Resolution

The invariant mass resolution, defined as the width σ of the Gaussian kernel of the Crystal Ball function, is estimated based on two independent procedures: the first method measures the mass resolution for the two charmonium resonances, J/ψ and $\psi(2S)$, as well as the three bottomonium resonances, $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$, decaying into $\mu^+\mu^-$. These five measurements are then used to interpolate the invariant mass resolution $\sigma_{\mu^+\mu^-}$ as a function of the invariant dimuon mass $m_{\mu^+\mu^-}$ to the central values of the B^0 and B_s^0 signals extracted in the previous section. This method is called the *Interpolation method*.
The second method using the decays $B^0_{(s)} \to h^+h'^-$ to estimate the invariant mass resolution is called the $B^0_{(s)} \to h^+h'^-$ *method*.

8.2.1 Determination from Charmonium and Bottomonium Resonances

Selection

The $c\bar{c}$ and $b\bar{b}$ resonance candidates for the interpolation method are selected in a similar way as the $B^0_{(s)} \to \mu^+\mu^-$ candidates, except for the selection criteria applied explicitly to select detached B mesons. The detailed selection cuts are listed in Tab. 8.4. Besides the charmonium and bottomonium resonances also a sample of Z candidates is selected.
Data samples for 2011 and 2012 as well as samples of simulated candidates for all five onia resonances and the decay $Z \to \mu^+\mu^-$ are selected. The simulated candidates were produced in the 2012 MC production campaign. Furthermore, a sample of dimuon pairs from simulated Drell-Yan candidates (2011 MC production campaign) have been selected with the same selection criteria as for the onia, except for the invariant mass cut. The Drell-Yan sample has the requirement of $m_{\mu^+\mu^-} > 2\,\text{GeV}/c^2$.

Simulation Studies

The invariant dimuon mass distributions of the dimuon resonances are described in the data samples by double Crystal Ball functions (${}_d\mathcal{CB}$). To prevent a possible bias from the correlation between the tail parameters $\alpha_{l,r}$, $n_{l,r}$ and the exponential shape of the distribution describing the combinatorial background, the tail parameters are extracted from MC and fixed in the fit to the data samples.
Figures 8.7 and 8.8 show the invariant mass distribution of the selected onia candidates in MC together with the double Crystal Ball function fitted using the binned maximum-likelihood method.

Table 8.4: Selection cuts for onia and $Z \to \mu^+\mu^-$ candidates used in the interpolation method.

Variable	**Onia**	$\boldsymbol{Z \to \mu^+\mu^-}$
Muons		
p_{T} [GeV/c]	> 1	> 20
p [GeV/c]	< 500	< 500
η	$[2.0, 4.5]$	$[2.0, 4.5]$
χ^2_{trk}/ndf	< 3	< 3
Ghost probability	< 0.3	< 0.3
IsMuon[a]	yes	yes
$\Delta\mathrm{LL}_{\mu-\pi}$	> -5	> -5
$\Delta\mathrm{LL}_{K-\pi}$	< 10	< 10
Onia/Z candidate		
χ^2_{vtx}/ndf	< 9	< 50
DOCA[b] [mm]	< 0.3	–
$m_{\mu^+\mu^-}$ [GeV/c^2]	$[3.01, 3.17]$ (J/ψ) $[3.60, 3.76]$ ($\psi(2S)$) $[9.00, 11.00]$ (Υ)	$[50, 130]$

[a] Muon identification as described in Sec. 3.4.1

[b] Distance of closest approach between the two muons

Also the tail parameters for $Z \to \mu^+\mu^-$ candidates are fixed from the corresponding MC sample. The invariant mass distribution in signal is shown Fig. 8.9. As the intrinsic decay width, Γ, of the Z boson is of the same order as the mass resolution, the distribution is fit by a convolution of a ${}_d\mathcal{CB}$ function and a Breit-Wigner distribution. The width parameter of the Breit-Wigner distribution is fixed to the nominal Z decay width $\Gamma = 2.495\,\mathrm{GeV}/c^2$ [81].
The resulting parameters of the onia and Z resonances are listed in Tab. 8.5. The tail parameters show an expected behaviour: first, they are very similar across the onia resonances, and second, the values for n_l are larger than for the hadronic B decays (*cf.* Tab. 7.1) as muons have on average less interaction with the detector material than charged hadrons. The same effect also enhances n_r with respect to the values observed in the $B^0 \to K^\pm\pi^\mp$ simulation, which indicates that the contribution of non-Gaussian resolution effects for onia candidates is smaller.

Invariant Dimuon Mass and its Resolution

A basic prerequisite for the interpolation method is that the behaviour of the invariant mass resolution as a function of the invariant mass, $\sigma_{\mu^+\mu^-}(m_{\mu^+\mu^-})$, can be predicted. Two ansatzes are used to extract this prediction. The first method uses the uncertainties on the track momenta and propagates them to the uncertainty on the invariant dimuon mass, while the second method uses directly the uncertainty on the invariant dimuon mass extracted from simulated Drell-Yan events.

The Modified Gluckstern Formula The Gluckstern Formula is a description of the relative uncertainty σ_p/p on the momentum measured by the bending of the trajectory of a charged

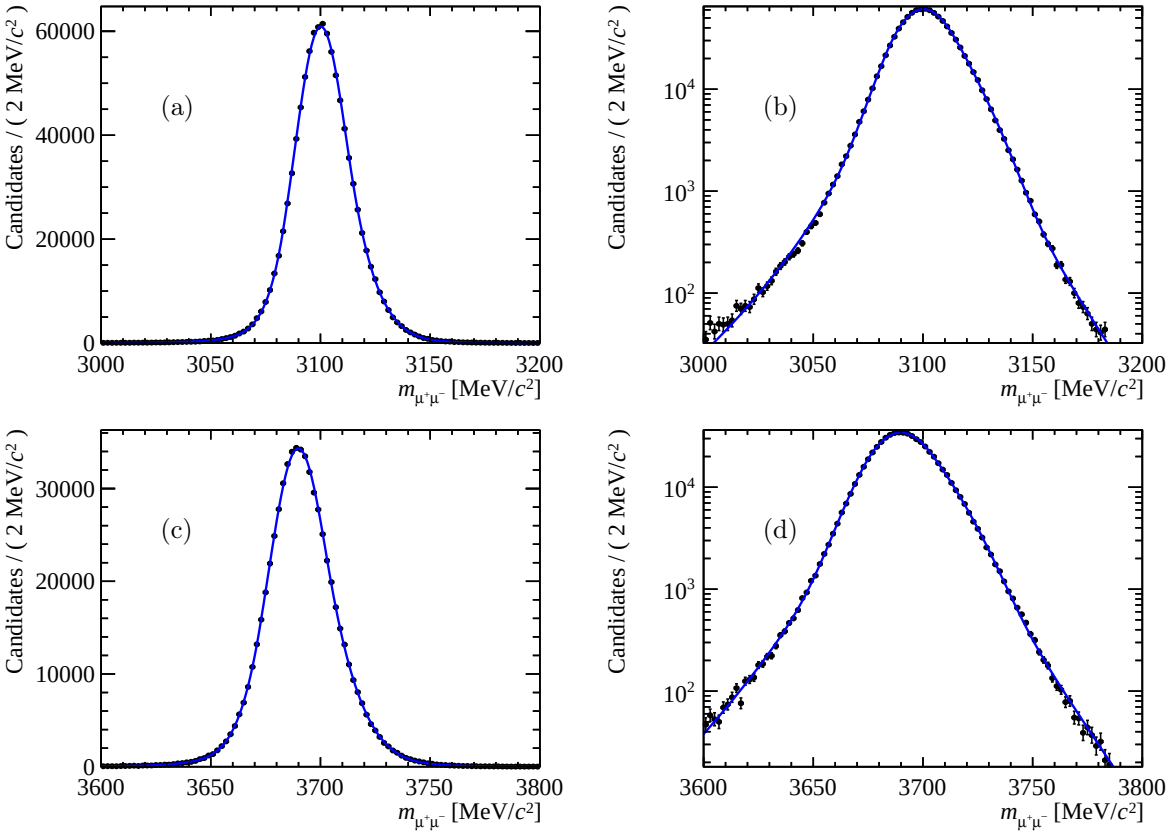


Figure 8.7: Invariant dimuon mass distribution in MC for (a) the J/ψ and (c) $\psi(2S)$ resonances. Figures (b) and (d) show the same distributions in log-scale. The blue lines show the fitted double Crystal Ball function.

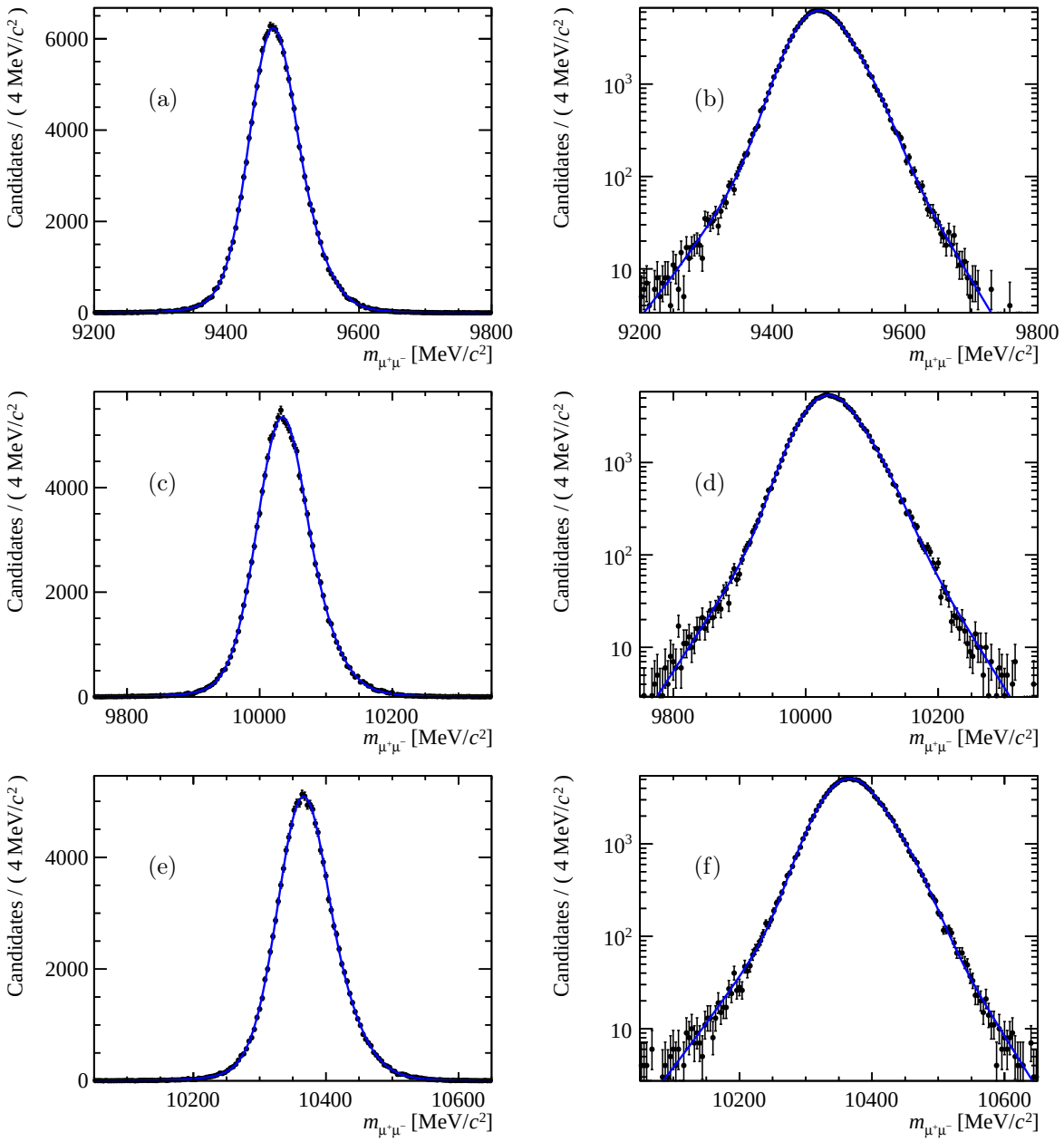


Figure 8.8: Invariant dimuon mass distribution in MC for (a) the $\Upsilon(1S)$, (c) $\Upsilon(2S)$ and (e) $\Upsilon(3S)$ resonances. Figures (b), (d) and (f) show the same distributions in log-scale. The blue lines show the fitted double Crystal Ball function.

Table 8.5: Mass parameters for the invariant mass distributions of $J/\psi \to \mu^+\mu^-$, $\psi(2S) \to \mu^+\mu^-$, $\Upsilon(1S) \to \mu^+\mu^-$, $\Upsilon(2S) \to \mu^+\mu^-$, $\Upsilon(3S) \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-$ in Monte Carlo extracted from maximum likelihood fits. The quoted uncertainties are those returned by the fit.

Parameter	$\boldsymbol{J/\psi \to \mu^+\mu^-}$	$\boldsymbol{\psi(2S) \to \mu^+\mu^-}$	
μ [MeV/c^2]	3100.002± 0.004	3689.870± 0.005	
σ [MeV/c^2]	12.893± 0.003	14.249± 0.004	
n_l	7.449± 0.004	7.449± 0.004	
n_r	81.286± 0.012	140.224± 0.016	
α_l	1.414± 0.008	1.529± 0.008	
α_r	-1.108± 0.008	-1.081± 0.007	
f	0.563± 0.005	0.563± 0.005	
Parameter	$\boldsymbol{\Upsilon(1S) \to \mu^+\mu^-}$	$\boldsymbol{\Upsilon(2S) \to \mu^+\mu^-}$	$\boldsymbol{\Upsilon(3S) \to \mu^+\mu^-}$
μ [MeV/c^2]	9470.301± 0.023	10032.234± 0.034	10366.352± 0.035
σ [MeV/c^2]	37.957± 0.018	39.968± 0.026	41.341± 0.028
n_l	6.837± 0.010	8.322± 0.013	7.580± 0.010
n_r	134.891± 0.011	133.662± 0.010	124.432± 0.015
α_l	1.529± 0.008	1.499± 0.010	1.504± 0.029
α_r	-1.081± 0.007	-1.086± 0.009	-1.112± 0.028
f	0.553± 0.005	0.523± 0.006	0.522± 0.013
Parameter	$\boldsymbol{Z \to \mu^+\mu^-}$		
μ [MeV/c^2]	91466 ±11		
σ [MeV/c^2]	1015 ±23		
n_l	1.313± 0.037		
n_r	15 ± 5		
α_l	1.04 ± 0.09		
α_r	-1.000± 0.017		
f	0.57 ± 0.05		

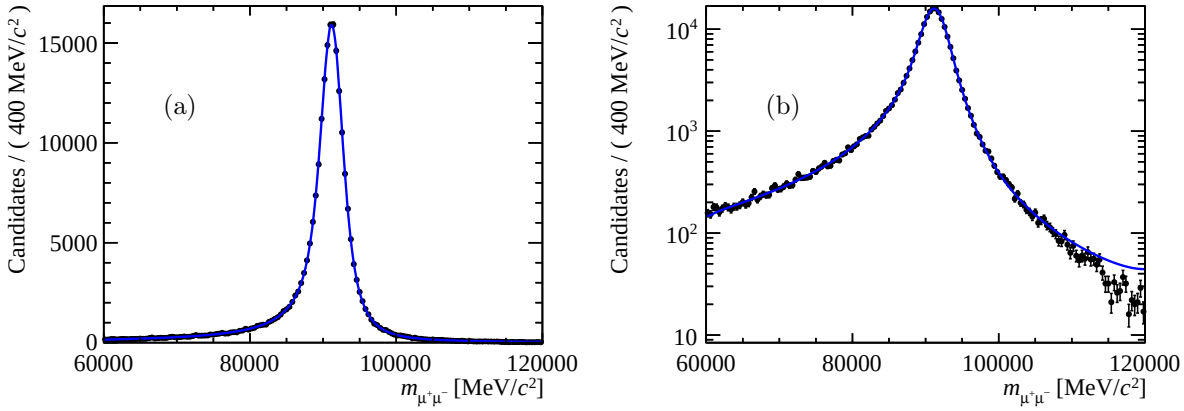


Figure 8.9: Invariant dimuon mass distribution in MC for the Z resonance in (a) linear and (b) log-scale. The blue lines show the fitted convolution of a double Crystal Ball function with a Breit-Wigner function.

particle in a homogenous magnetic field [159,179]. This formula predicts that σ_p/p is the square-root of the sum in quadrature of a term linear in momentum p, a term proportional to p^{-1} and a constant term

$$\frac{\sigma_p}{p} = \sqrt{(A \cdot p)^2 + (B \cdot p^{-1})^2 + C^2}. \tag{8.1}$$

The linear term represents the resolution due to the finite resolution of the tracking detectors, the second term the resolution component from multiple scattering and the third term the resolution component from the uncertainty on the angle of the track with respect to the magnetic field lines.
This formula is only exact in an ideal environment, *e.g.* constant magnetic field strength over the measured trajectory, equidistant arrangement of the tracking stations and independence of the number of measured points on the track momentum.
All these assumptions do not hold for the LHCb detector: the tracking detectors are outside of the magnetic field, there is no uniform distribution of the tracking stations along the z-axis of the experiment and also the number of hits measured in the different tracking sub-detectors is not momentum-independent as shown in Ref. [137]. Further the different tracking detectors have also different position resolutions.

The behaviour of σ_p/p as a function of p for muons in LHCb is studied using dimuon pairs from simulated Drell-Yan events with m_{γ^*} above $2\,\text{GeV}/c^2$. The resolution σ_{p_μ} is estimated as the standard deviation of the difference between the reconstructed and the true momenta of the muons. Figure 8.10 shows σ_{p_μ}/p_μ as a function of p_μ extracted from this sample. A fit of the Gluckstern formula to the data shown by the long-dashed blue line fails as expected due to the not-fulfilled assumptions. All issues mentioned above actually affect the contribution due to the resolution of the tracking detectors. A possible modification to take these issues into account is the introduction of an additional parameter $\gamma > 1$ describing how far the first term deviates from its linear behaviour. This modified Gluckstern formula is

$$\frac{\sigma_p}{p} = \sqrt{(A \cdot p^{\gamma-1})^2 + (B \cdot p^{-1})^2 + C^2}. \tag{8.2}$$

Figure 8.10 also shows the fit of this formula to the data by the solid red line. The individual contributions from the detector resolution and multiple scattering components are shown by the red dotted and red dashed lines, respectively. It can be concluded that this formula describes the momentum resolution well in a wide range from 2 to 200 GeV/c. The obtained χ^2/ndf from the fit of the modified Gluckstern formula is 321.54/156. The extracted value of the exponent is $\gamma = 1.3862 \pm 0.0014$.

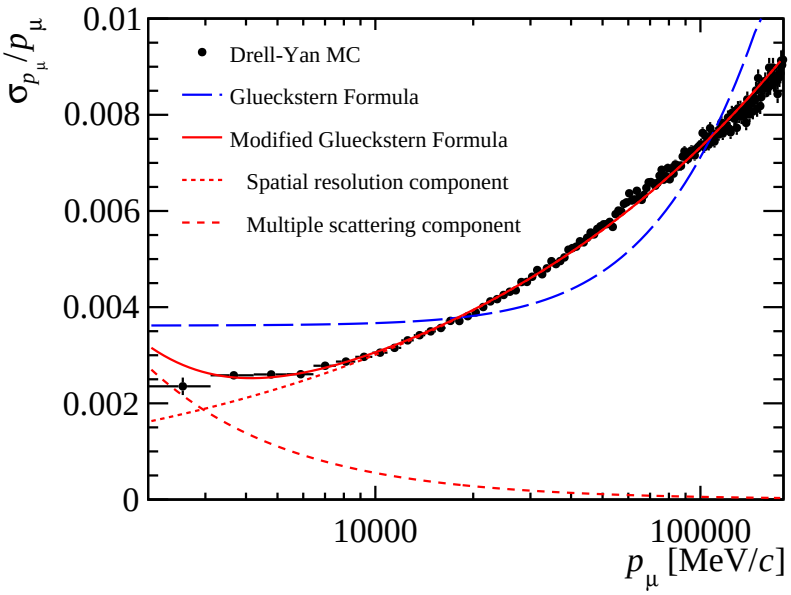


Figure 8.10: Relative momentum resolution σ_{p_μ}/p_μ as a function of the momentum p_μ of muons from Drell-Yan MC. The blue long-dashed line shows the fitted Gluckstern formula while the red solid line shows the modified Gluckstern formula defined in Eq. (8.2). The red dashed and dotted lines show the multiple scattering and tracking detector resolution components, respectively. The constant component associated to the angular resolution is negligible due to the – on average – large angle between the magnetic field lines and the tracks.

Figure 8.10 also shows that the detector resolution dominates for momenta above 3 GeV/c. Thus, the momentum resolution can be described as

$$\sigma_{p_\mu} \propto p_\mu^\gamma, \tag{8.3}$$

for most of the considered track momentum range. The invariant dimuon mass $m_{\mu^+\mu^-}$ is

$$m^2_{\mu^+\mu^-} = 2(E_{\mu^+} \cdot E_{\mu^-} - \vec{p}_{\mu^+} \cdot \vec{p}_{\mu^-}) + 2m^2_\mu \tag{8.4}$$

and simplified using $|\vec{p}_{\mu^\pm}| \gg m_\mu$ as well as $m_{\mu^+\mu^-} \gg m_\mu$

$$m^2_{\mu^+\mu^-} = 2p_{\mu^+}p_{\mu^-}(1 - \cos\theta), \tag{8.5}$$

where $p_{\mu^\pm} = |\vec{p}_{\mu^\pm}|$ and θ is the angle between the two muons.
The uncertainty on $m^2_{\mu^+\mu^-}$, neglecting correlations among the three variables $p_{\mu^\pm}$ and θ, can be calculated from error propagation as

$$\sigma^2_{m^2_{\mu^+\mu^-}} = m^4_{\mu^+\mu^-} \cdot \left[\left(\frac{\sigma_{p_{\mu^+}}}{p_{\mu^+}} \right)^2 + \left(\frac{\sigma_{p_{\mu^-}}}{p_{\mu^-}} \right)^2 \right] + (2p_{\mu^+}p_{\mu^-}\sigma_{\cos\theta})^2. \tag{8.6}$$

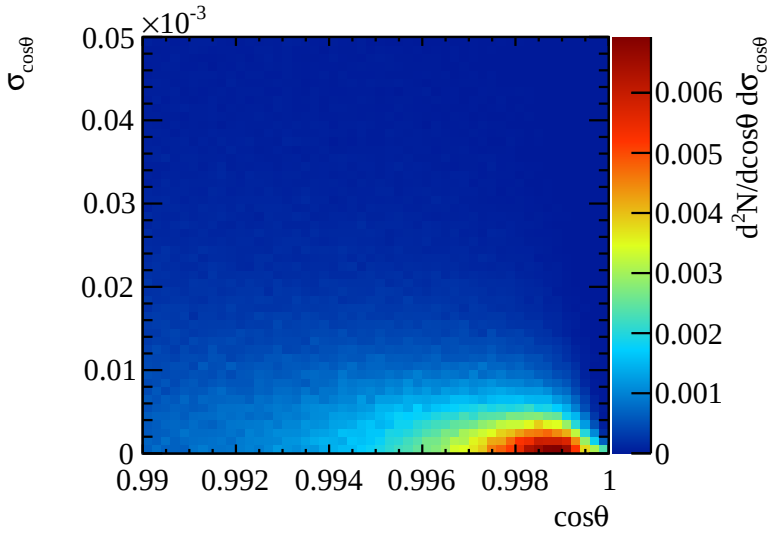


Figure 8.11: Two-dimensional distribution of the cosine of the angle θ between the positive and negative muon in Drell-Yan MC and the uncertainty on the cosine of the angle. As $\cos\theta$ is on average large and its uncertainty small, the impact of the uncertainty of $\cos\theta$ on the uncertainty of the invariant dimuon mass is negligible.

Figure 8.11 shows the two-dimensional distribution of $\cos\theta$ and the uncertainty on $\cos\theta$, $\sigma_{\cos\theta}$, of the muons in the Drell-Yan MC sample. The latter is taken as the standard deviation of the difference between the true and reconstructed value of θ.
This distribution shows that the last term in Eq. (8.6) can be neglected as $\sigma_{\cos\theta}$ is small.
Error propagation allows to translate the uncertainty on $m^2_{\mu^+\mu^-}$ into one on $m_{\mu^+\mu^-}$ by

$$\sigma_{\mu^+\mu^-} = \frac{\sigma_{m^2_{\mu^+\mu^-}}}{2m_{\mu^+\mu^-}} \tag{8.7}$$

$$= \frac{m_{\mu^+\mu^-}}{2} \cdot \sqrt{\left(\frac{\sigma_{p_{\mu^+}}}{p_{\mu^+}}\right)^2 + \left(\frac{\sigma_{p_{\mu^-}}}{p_{\mu^-}}\right)^2} \tag{8.8}$$

and with Eq. (8.3)

$$\propto \frac{1}{\sqrt{2}} m_{\mu^+\mu^-} p^{\gamma-1}_{\mu^\pm}. \tag{8.9}$$

Equation (8.5) shows that approximately $m_{\mu^+\mu^-} \propto p_{\mu^\pm}$. Thus, the invariant mass resolution as a function of the invariant dimuon mass can be described by

$$\sigma_{\mu^+\mu^-}(m_{\mu^+\mu^-}) = a_0 m^{\gamma}_{\mu^+\mu^-}. \tag{8.10}$$

To incorporate possible non-mass-dependent resolution effects, which have been neglected and can dominate close to $m_{\mu^+\mu^-} = 0$ (*e.g.* from multiple scattering), an additional constant term is added, such that the final power-law like function, which is used to describe $\sigma_{\mu^+\mu^-}(m_{\mu^+\mu^-})$ is

$$\sigma_{\mu^+\mu^-}(m_{\mu^+\mu^-}) = a_0 m^{\gamma}_{\mu^+\mu^-} + a_1. \tag{8.11}$$

Drell-Yan Monte Carlo The second ansatz directly uses the invariant mass of the dimuon candidates in the selected Drell-Yan MC sample. The invariant mass resolution is calculated in bins of $m_{\mu^+\mu^-}$ as the standard deviation of the distribution of the differences between the true and the reconstructed invariant dimuon mass.
Figure 8.12 shows the extracted resolution as a function of $m_{\mu^+\mu^-}$. The data points are fit with the function in Eq. (8.11), in which all three parameters $a_{0,1}$, γ are free, and with a linear function in the range 3 to 11 GeV/c^2.

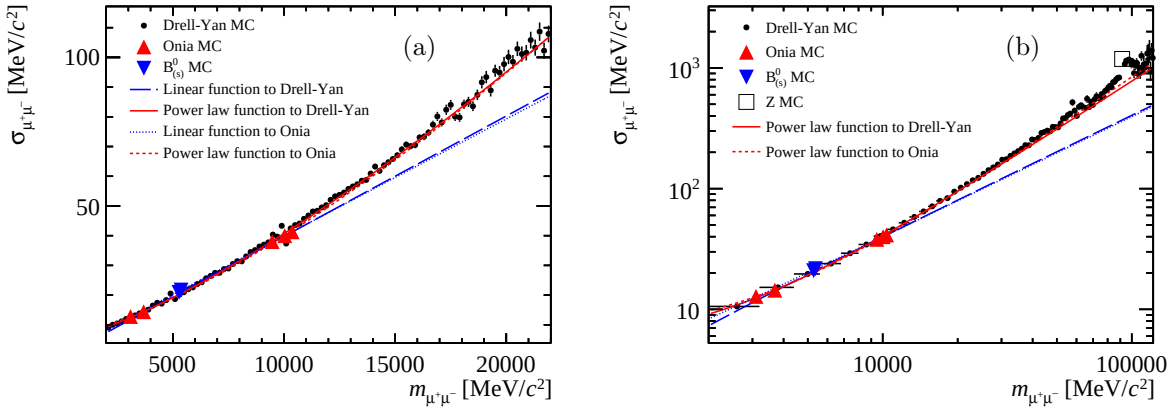


Figure 8.12: Invariant dimuon mass resolution as a function of the invariant dimuon mass in Drell-Yan MC: (a) mass range from 2 to 22 GeV/c^2, (b) mass range from 2 to 120 GeV/c^2. The values extracted for the onia and Z resonances from MC as well as for two-body decays of $B^0_{(s)}$ from MC are also shown in the plot. The red lines show the fit of the power-law like function described in Eq. (8.11) to the Drell-Yan data (solid line) and to the five onia resonances (dashed line). The blue lines show the fit of a linear function to the Drell-Yan data (long-dashed line) and to the five onia resonances (dotted line).

While the fit quality of the linear function is poor giving $\chi^2/\text{ndf} = 284.2/38$, it is reasonable good given a value of $\chi^2/\text{ndf} = 61.3/37$ for the power-law like function. The extracted exponent $\gamma = 1.33 \pm 0.07$ is in good agreement with the exponent extracted from the fit of the relative momentum resolution.
Figure 8.12 also shows the invariant mass resolution in MC of the charmonium and bottomonium resonances and of the Z resonance (*cf.* Tab. 8.5). In addition, also the resolutions for $B^0_s \to \mu^+\mu^-$ and $B^0 \to K^\pm\pi^\mp$ in MC, listed in Tab. 7.1, are shown.
The data points from the onia resonances are fit as well by the function described in Eq. (8.11) and by a linear function. Also in this case the power-law like function gives a better fit in terms of χ^2/ndf (3.6/2) than the linear one (10.8/3), and the extracted exponent $\gamma = 1.39 \pm 0.20$ is in agreement with the values extracted by the other method. Table 8.6 shows the interpolated values of the mass resolution for $B^0_{(s)}$. The predictions from the power-law like function are closer to the values extracted directly from the $B^0_{(s)}$ MC samples than the prediction based on the linear function.
In summary, the functional form proposed in Eq. (8.11) describes the behaviour of the invariant mass resolution in the relevant region $[3, 11]$ GeV/c^2 well. The predictions based on the power-law like function underestimates the resolution for very high dimuon masses (*i.e.* $m_{\mu^+\mu^-} = M_Z$) as shown in Fig. 8.12(b).

Table 8.6: Interpolated mass resolution for the $B^0_{(s)}$ masses in MC based on charmonium and bottomonium resonances using a linear and a power-law like interpolation model. The last row shows for comparison the mass resolution extracted from a fit of the mass distribution in $B^0_s \to \mu^+\mu^-$ (B^0_s) and $B^0 \to K^\pm\pi^\mp$ (B^0), respectively in MC.

Particle	Linear Function	Power-law like Function	MC fit
$\sigma_{\mu^+\mu^-}(m_{B^0})$ (σ_{B^0}) [MeV/c^2]	21.25	20.43	20.61
$\sigma_{\mu^+\mu^-}(m_{B^0_s})$ ($\sigma_{B^0_s}$) [MeV/c^2]	21.65	20.86	20.91

Fit Model and Results

The invariant dimuon mass distributions of the onia samples selected in data as described above are shown in Figs 8.13 and 8.14 for 2011 and for 2012 data, respectively. Each resonance is modelled by a ${}_d\mathcal{CB}$ function where the tail parameters $\alpha_{l,r}$ and $n_{l,r}$ as well as the function f are fixed to the values determined in MC. The combinatorial background is described in each of the three invariant mass ranges (*i.e.* J/ψ, $\psi(2S)$ and Υ sample) by an individual exponential distribution.
The fit to extract the invariant mass resolution of the onia resonances is based on the binned maximum-likelihood method. The χ^2 of the fit is calculated as

$$\chi^2 = \sum_i \frac{(n_i - N \cdot f(m_i))^2}{\sigma_i^2}, \tag{8.12}$$

where index i runs over the number of mass bins; n_i is the number of candidates in the i-th bin, σ_i the statistical uncertainty on n_i, $f(m_i)$ the value of the fitted distribution normalised to the fit range evaluated at the bin center and N the total number of candidates in the fit range. The number of degrees of freedom is defined in the usual way, *i.e.* the number of bins minus the number of free fit parameters.
The obtained values for the χ^2/ndf are high as the fit model does not describe the resonance well. A modification of the fit model is not feasible as it has to be the same model that is used for the description of the invariant dimuon mass distribution in $B^0_{(s)} \to \mu^+\mu^-$. Otherwise the meaning of the estimated parameters changes.
The uncertainties extracted from the fit for the free parameters are scaled by the factor $\sqrt{\chi^2/\text{ndf}}$. The pull as a function of the invariant mass based on the distribution with the modified parameter uncertainties is shown by the grey band at the bottom of the plots in Figs 8.13 and 8.14. There is a residual imperfection of the fit, which is visible in the tail regions of the resonances due to the non-optimal fit model.
The same procedure is also performed for the Z samples in 2011 and 2012 data, shown in Fig. 8.15. In this case the signal is described by a convolution of the ${}_d\mathcal{CB}$ function and a Breit-Wigner distribution whose width is fixed to the nominal value of the width of the Z resonance. The background and the Drell-Yan component are again described by an exponential distribution.
Table 8.7 lists the extracted values for the resolution as well as the central value of the resonances given by the parameter μ of the ${}_d\mathcal{CB}$ function. For the onia resonances also the ratios of the signal yield of a certain resonance to the total charmonium or bottomonium signal yields are

listed.

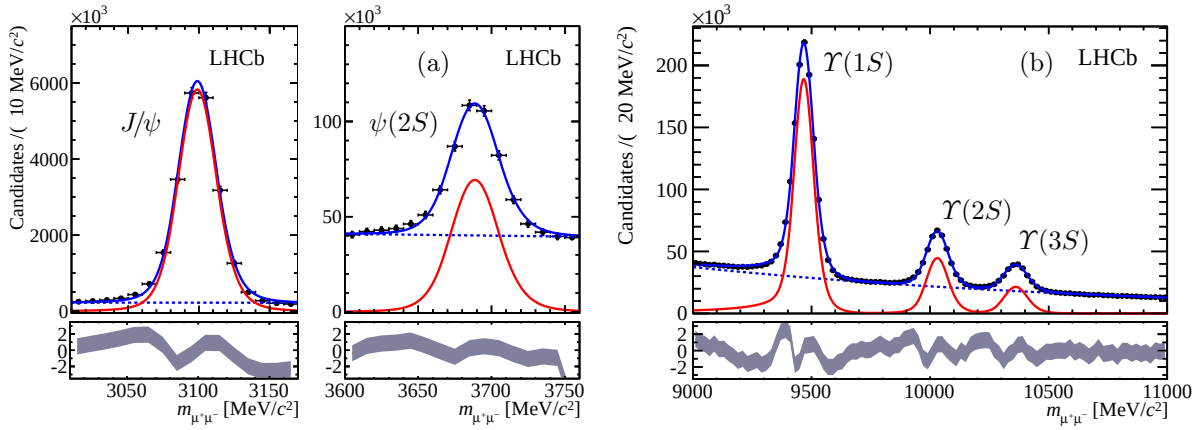


Figure 8.13: Invariant dimuon mass distribution of (a) charmonium resonances (left: J/ψ, right: $\psi(2S)$) and (b) bottomonium resonances in 2011 data. The distributions are fit with the sum of two Crystal Ball functions with common central value and width of the Gaussian kernel for each resonance (red solid lines) and an exponential distribution (blue dashed line) for the combinatorial background. The total fit model is shown by the blue solid line.

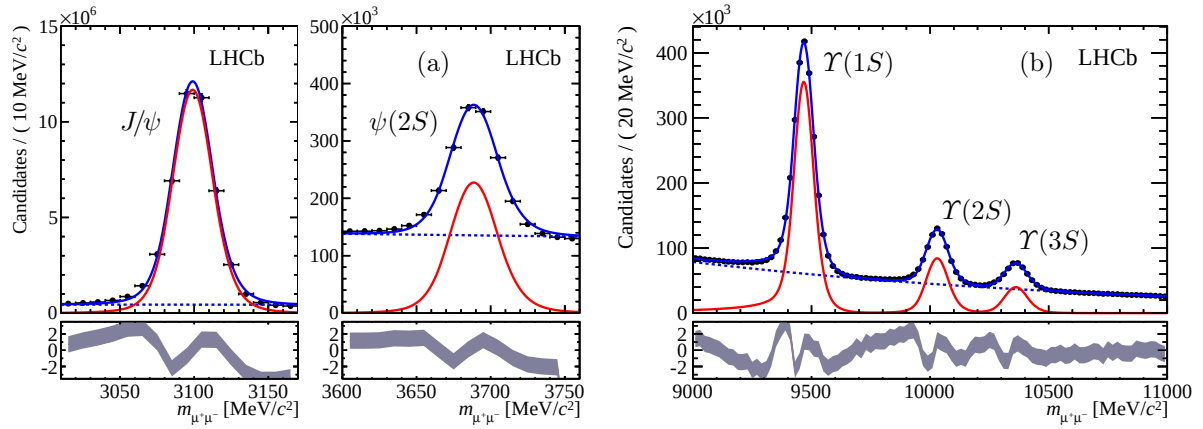


Figure 8.14: Invariant dimuon mass distribution of (a) charmonium resonances (left: J/ψ, right: $\psi(2S)$) and (b) bottomonium resonances in 2012 data. The distributions are fit with the sum of two Crystal Ball functions with common central value and width of the Gaussian kernel for each resonance (red solid lines) and an exponential distribution (blue dashed line) for the combinatorial background. The total fit model is shown by the blue solid line.

The central values μ and the resolution parameters σ are used to perform an interpolation of the invariant mass resolution based on Eq. (8.11). The results of the fits are shown in Fig. 8.16. The fit is performed without constraints on the three parameters a_0, a_1 and γ. The obtained fit quality given by χ^2/ndf is 3.1/2 for 2011 and 3.3/2 for 2012 data. The corresponding numbers for a linear fit are 7.3/3 and 7.1/3, respectively.
The extracted exponent parameters are $\gamma = 1.25 \pm 0.13$ for 2011 data and $\gamma = 1.25 \pm 0.09$ for 2012 data. These values are in good agreement with the expectations from the above described

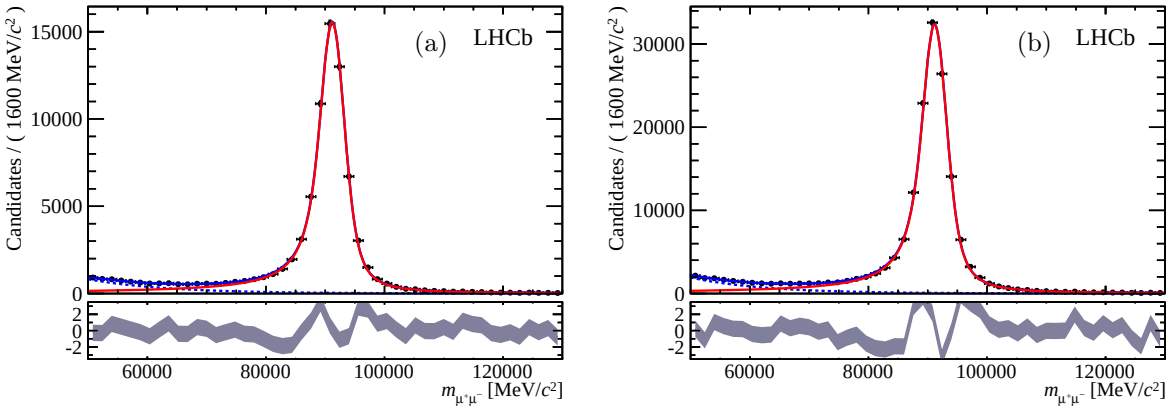


Figure 8.15: Invariant dimuon mass distribution of Z candidates in (a) 2011 and (b) 2012 data. The distributions are fit for the signal component with the convolution of the sum of two Crystal Ball functions with common central value and the width of the Gaussian kernel and a Breit-Wigner distribution whose width is fixed to the nominal intrinsic width of the Z resonance (red solid line). The background and the Drell-Yan component are described by an exponential distribution (blue dashed line).

Table 8.7: The mass resolution σ and the central value μ for $J/\psi \to \mu^+\mu^-$, $\psi(2S) \to \mu^+\mu^-$, $\Upsilon(1S) \to \mu^+\mu^-$, $\Upsilon(2S) \to \mu^+\mu^-$, $\Upsilon(3S) \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-$ in 2011 and 2012 data extracted from maximum likelihood fit methods to the invariant dimuon mass. The last row shows the fractions of the individual resonance yields to the sum of the yields of the charmonium or bottomonium signal, respectively.

Resonance	$\boldsymbol{\mu}$ **[MeV/c^2]**	$\boldsymbol{\sigma}$ **[MeV/c^2]**	**fraction** $\boldsymbol{f}$
2011 Data			
J/ψ	3098.910± 0.023	13.467± 0.029	0.962±0.003
$\psi(2S)$	3688.63 ± 0.18	15.74 ± 0.24	0.038±0.003
$\Upsilon(1S)$	9467.26 ± 0.16	44.91 ± 0.21	0.733±0.006
$\Upsilon(2S)$	10030.2 ± 0.3	46.89 ± 0.5	0.177±0.005
$\Upsilon(3S)$	10361.1 ± 0.5	48.75 ± 0.7	0.090±0.008
Z	91330 ±48	1531 ±29	–
2012 Data			
J/ψ	3098.970± 0.009	13.466± 0.013	0.961±0.003
$\psi(2S)$	3688.59 ± 0.11	15.80 ± 0.14	0.039±0.003
$\Upsilon(1S)$	9467.23 ± 0.08	44.51 ± 0.15	0.737±0.004
$\Upsilon(2S)$	10029.70 ± 0.16	46.82 ± 0.27	0.177±0.003
$\Upsilon(3S)$	10361.70 ± 0.28	49.52 ± 0.5	0.086±0.005
Z	91290 ±40	1500 ±21	–

study using the modified Gluckstern formula and Drell-Yan MC.
The interpolated values for the expected mass resolution parameters at the masses of the B^0 and B_s^0 mesons are listed in Tab. 8.8 where the central values extracted in Sec. 8.1 are used as mass values for $B_{(s)}^0$.

Table 8.8: Dimuon mass resolution parameter σ for B^0 and B_s^0 in 2011 and 2012 data based on the interpolation method.

Data sample	**Resolution parameter σ [MeV/c^2]**	
	$\boldsymbol{B^0}$	$\boldsymbol{B_s^0}$
2011 Data	23.11±0.22	23.52±0.23
2012 Data	23.21±0.16	23.63±0.17

As cross check the result of the fit to the five onia resonances is extrapolated to the Z mass as shown in Fig. 8.17. The predicted invariant dimuon mass resolution is $645 \pm 8\,\mathrm{MeV}/c^2$ from 2011 data and $651 \pm 13\,\mathrm{MeV}/c^2$ from 2012 data to be compared to the measured values. The prediction of the invariant dimuon mass resolution fails, as expected from the MC studies. The larger deviation in data compared to MC is due to additional reconstruction effects in data such as an observed dependence of the Z mass on the polar angle of the muons [190]. Therefore the observed Z mass resolution in data is about 50 % worse than in MC. As the polar angle dependency increases with the momentum of the muons, this effect is much smaller for resonances or particles with a lower mass, *e.g.* charmonium and bottomonium resonances as well as *e.g.* B^0 and B_s^0 mesons. A reason is that the alignment of the tracking stations uses $D^0 \to K^\pm\pi^\mp$ and $J/\psi \to \mu^+\mu^-$ candidates.

Systematic Uncertainties

Four sources of systematic uncertainties on the interpolated mass resolution are taken into account:

1. **Selection criteria** The effect of the selection criteria are estimated by relaxing the cut on the muon p_T to $0.75\,\mathrm{GeV}/c$ and removing the cut on the muon η, as these two criteria are not identical to the ones in the selection of the $B_{(s)}^0 \to \mu^+\mu^-$ candidates. The interpolation is repeated and the difference to the nominal resolution is assigned as systematic uncertainty.

2. **Interpolation function** The uncertainty based on the choice of the interpolation function is taken as the difference between the interpolated mass resolution in MC and the mass resolution extracted from the fit of the $B_s^0 \to \mu^+\mu^-$ invariant mass distribution in MC. For B^0, the same systematic uncertainty as for B_s^0 is assigned.

3. **Uncertainty asymmetry** The different yields of the onia resonances lead to large differences in the statistical uncertainties on the extracted mass resolutions. Therefore, the interpolation is dominated by the J/ψ and the $\Upsilon(1S)$ resonances, but as the fit function has three free parameters, statistical fluctuations in the determination of the resolution for the other resonances can significantly affect the curvature of the function (*i.e.* the parameter γ). The interpolation is repeated with the uncertainties on the mass resolutions in all five

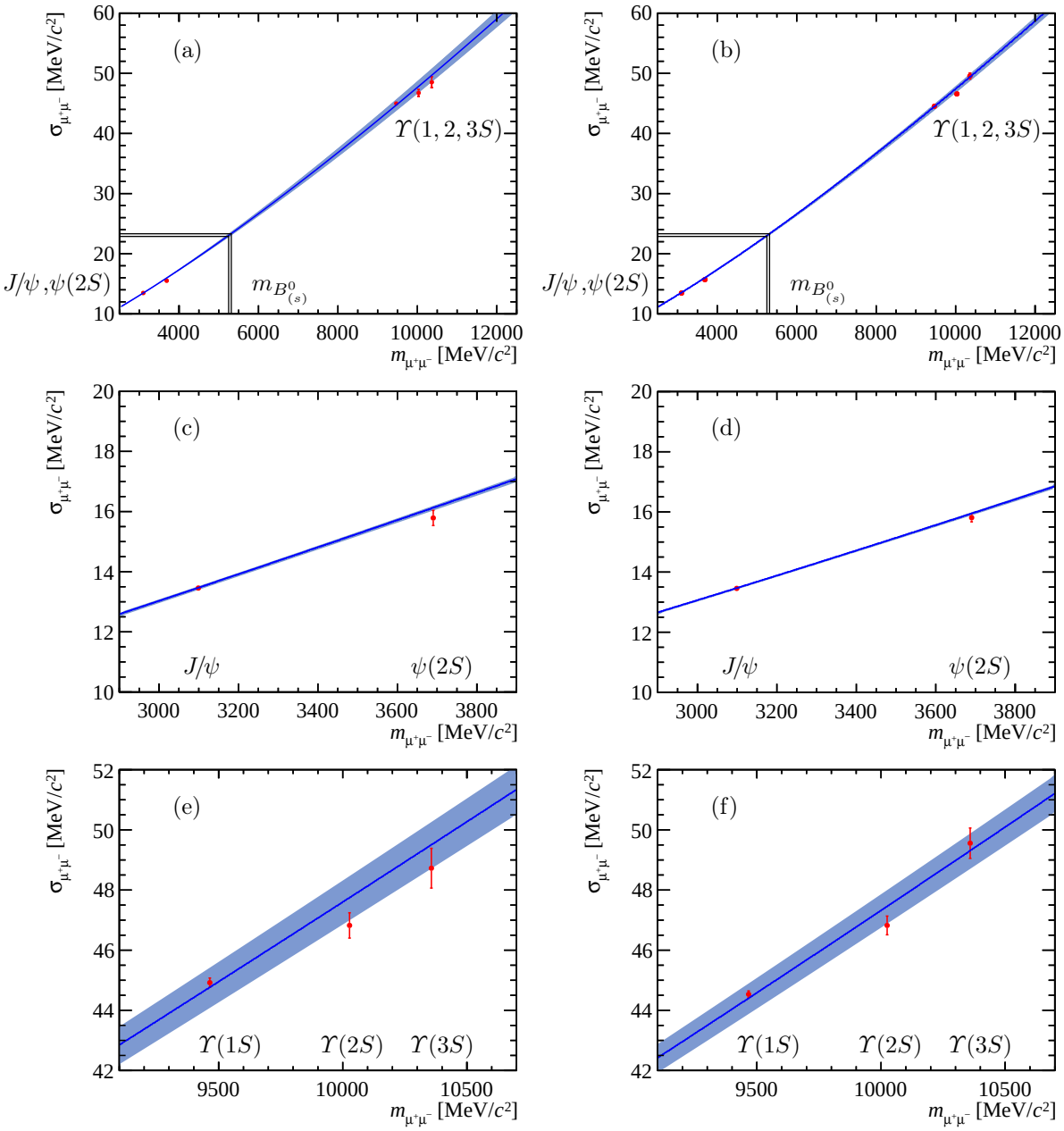


Figure 8.16: Interpolation of the invariant dimuon mass resolution for B^0_s and B^0 in (a) 2011 and (b) 2012 data based on Eq. (8.11) using the measured resolution of the J/ψ and $\psi(2S)$ as well as the $\Upsilon(nS)$ resonances. The blue line shows the fitted function while the blue shaded region shows the statistical uncertainty on the interpolation. Figures (c) and (d) show the zoom into the charmonium region and Figs (e) and (f) into the bottomonium region.

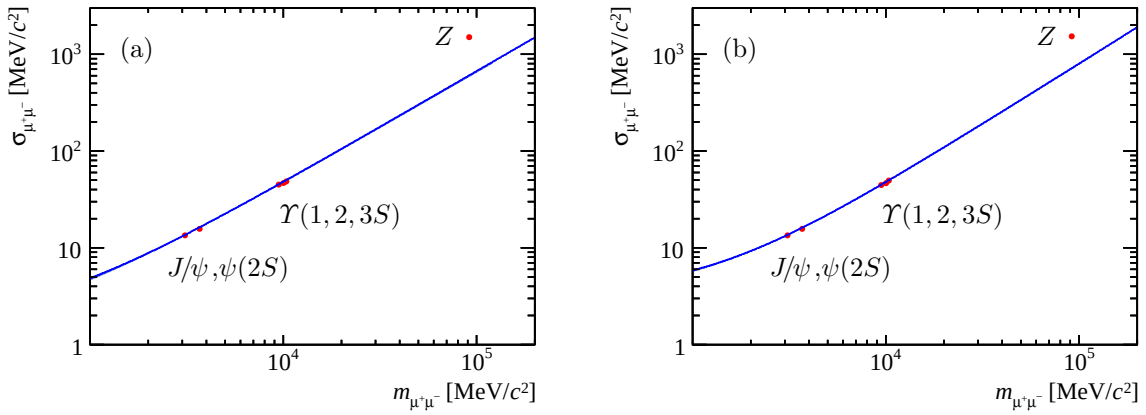


Figure 8.17: Extrapolation of the invariant dimuon mass resolution in (a) 2011 and (b) 2012 data based on Eq. (8.11) using the measured resolution of the J/ψ and $\psi(2S)$ as well as the Υ resonances to the Z mass. The blue line shows the fitted function while the blue shaded region shows the statistical uncertainty on the interpolation.

resonances set to the same value and the difference of the obtained interpolated resolutions for $B^0_{(s)}$ with respect to the original values is assigned as systematic uncertainty.

4. **Mass window** The invariant mass resolution changes as a function of the invariant mass across the $B^0_{(s)}$ signal mass window. As a systematic uncertainty for this effect the full change in the value of the fitted power-law like function across the $\pm 60\,\text{MeV}/c^2$ signal mass window used in the extraction of upper limits of the branching fractions is assigned.

The values of all systematic uncertainties are listed in Tab. 8.9 as well as the total uncertainty calculated by adding the individual contributions in quadrature.
In addition, a systematic check of the invariant dimuon mass fit model has been performed. The invariant dimuon mass distribution in the sidebands of the onia resonances (defined as mass regions that are more than 40 (J/ψ), 50 ($\psi(2S)$) or 130 (Υ) MeV/c^2 away from the central value of the resonance) was fit by an exponential distribution describing the combinatorial background. The shape and yield of the combinatorial background are then fixed and a Gaussian distribution is fit to the mass distribution of each resonance in a region of $\pm 1\sigma$ around the central value of the resonance. The value of σ is the mass resolution determined in the nominal fit. This alternative fit procedure is independent of the tails of the resonances and allows therefore to test for the impact of the description of the tails based on MC. The extracted mass resolution for B^0 and B^0_s based on this alternative fit are within the statistical uncertainties compatible with the values extracted from the standard fit. Therefore no additional systematic uncertainty is assigned.
The final values for the width parameter σ determined using the interpolation method are

$$\begin{aligned}\sigma_{B^0} &= 23.11 \pm 0.22(\text{stat.}) \pm 0.41(\text{syst.})\,\text{MeV}/c^2\\ \sigma_{B^0_s} &= 23.52 \pm 0.23(\text{stat.}) \pm 0.48(\text{syst.})\,\text{MeV}/c^2\end{aligned}$$

for 2011 data and

$$\begin{aligned}\sigma_{B^0} &= 23.21 \pm 0.16(\text{stat.}) \pm 0.43(\text{syst.})\,\text{MeV}/c^2\\ \sigma_{B^0_s} &= 23.63 \pm 0.17(\text{stat.}) \pm 0.45(\text{syst.})\,\text{MeV}/c^2\end{aligned}$$

Table 8.9: Systematic uncertainties on the mass resolution estimated by the interpolation method.

Data Set	Systematic uncertainties [MeV/c^2]				
	Selection	**Interpolation funciton[a]**	**Uncertainty assymmetry**	**Mass window**	**Total**
2011 Data					
B^0	0.20	0.15	0.17	0.28	0.41
B_s^0	0.22	0.15	0.28	0.28	0.48
2012 Data					
B^0	0.19	0.15	0.21	0.28	0.43
B_s^0	0.23	0.15	0.23	0.28	0.45

[a] The same systematics for both data sets and also for B_s^0 as well as B^0 has been taken based on the difference between the result of the interpolation method and the fit of the invariant dimuon mass distribution of $B_s^0 \to \mu^+\mu^-$ in MC.

for 2012 data. The ratio β_{reso} between the mass resolutions for B^0 and B_s^0 is

$$\beta_{\text{reso}} = 1.0177 \pm 0.0010$$

for 2011 data and

$$\beta_{\text{reso}} = 1.0181 \pm 0.0006$$

for 2012 data. The quoted uncertainty is the combined statistical and systematic uncertainty, where it has been assumed that the mass resolution values for B^0 and B_s^0 are fully correlated. These ratios serve as input to the BDT calibration discussed in the previous chapter as well as to the mass resolution determination based on the decays $B_{(s)}^0 \to h^+h'^-$ discussed in the following section.

8.2.2 Determination from the Decays $B_{(s)}^0 \to h^+h'^-$

The estimation of the invariant mass resolution using the decays $B_{(s)}^0 \to h^+h'^-$ relies on similar methods as the BDT calibration for signal discussed in the Chapter 7.
An estimation of the invariant mass resolution with $B_{(s)}^0 \to h^+h'^-$ decays has the mis-identification of pions as kaons and vice versa as main obstacle. Candidates with one or both particles mis-identified lead to a broadening of the signal resonances in the dihadron spectrum. The mis-identification can be suppressed by applying PID-information, but the efficiency of cuts on the ΔLL variables discussed in Chapter 7 is momentum dependent and therefore affects the invariant mass resolution. This bias needs to be corrected for the determination of the invariant mass resolution.

Selection and Data Sample

The selection of the $B_{(s)}^0 \to h^+h'^-$ candidates is the same as in the sample used for the BDT calibration, except that no condition is applied here that the events have to be triggered independently of the $B_{(s)}^0 \to h^+h'^-$ candidate.

As in the BDT calibration, the variable $\kappa \geq 0$ is used to describe the cuts on the ΔLL variables and the width of the B^0 and B_s^0 resonances are studied as a function of this variable. For a given value of κ, a final state particle is identified as a kaon if $\Delta\mathrm{LL}_{K-\pi} > \kappa$ and as a pion if $\Delta\mathrm{LL}_{K-\pi} < -\kappa$. Candidates with $|\Delta\mathrm{LL}_{K-\pi}| < \kappa$ for at least one final state particle are rejected. To correct for the bias of these cuts, the accepted candidates are weighted in the same way as in the BDT calibration, which is defined in Eq. (7.8). As the PID efficiency is only calculated in the momentum range of 5 to 200 GeV/c, all events where one of the track momenta exceeds 200 GeV/c are rejected. The result is later corrected for this cut. The determination of the invariant mass resolution is performed separately for the 2011 and 2012 data samples.

Fit Model and Results

The invariant mass distributions are considered for values of κ in the range from 0 to 20 and in steps of 0.5. Figures 8.18 and 8.19 show the invariant mass distributions of the weighted candidates for different values of κ. The distributions are fit by the same fit model as for the BDT calibration. The ratio of the mass resolution parameters σ_{B^0} and $\sigma_{B_s^0}$ are fixed to the ratios β_{reso} determined in Sec. 8.2.1.
The mass resolution parameter of the Gaussian kernel describing the B^0 signal, σ_{B^0}, – in this context called σ_{raw} – as a function of κ is shown in Fig. 8.20. Besides a component of the width that is constant as a function of κ and can interpreted as the actual mass resolution, there is a component that decreases as a function of κ. This decreasing component is associated to mis-identified final state particles causing a broadening of the signal and should decrease linearly with the yield of mis-identified candidates. Due to the definition of the $\Delta\mathrm{LL}_{x-y}$ variables as the logarithm of the ratio of the probabilities that a particle is of type x or of type y, the component for the mis-identification of a single final state particle should have the form

$$\sigma_{\text{mis-id}}(\kappa) = \frac{a_0}{1 + \exp(\kappa)}. \tag{8.13}$$

The width, σ_{raw}, of the signal as a function of κ is described by the function

$$\begin{aligned} \sigma_{\mathrm{raw}}(\kappa) &= \sigma_{\text{mass-reso}} + \sigma'_{\text{mis-id}}(\kappa) \\ &= \sigma_{\text{mass-reso}} + \frac{a_0}{1 + r \cdot \exp(\kappa/\kappa_0)}, \end{aligned} \tag{8.14}$$

where the parameters r and κ_0 have been introduced to take into account the deviation of the $\Delta\mathrm{LL}_{x-y}$ variable from its nominal meaning described in Eq. (8.13) due to imperfections in the calibration of the ΔLL variables.
The fit of Eq. (8.14) to 2011 data and to 2012 data is shown in Fig. 8.20, where the fit range in 2011 data is $0 < \kappa < 13$ and in 2012 data $0 < \kappa < 17$. The reason for the different fit ranges is the drop in σ_{raw} in the 2011 data at about $\kappa = 15$. This effect is caused by limited statistics in the high momentum region. As shown in Fig. 7.7, candidates including high momentum final state particles are almost completely rejected by tight cuts on the $\Delta\mathrm{LL}_{K-\pi}$ variable. Therefore, above a certain value of κ a significant number of high momentum bins in the three-dimensional binning (p, η and number of tracks) used for the PID efficiency correction are empty and thus the reweighting of the events according to the PID efficiency cannot reproduce the unbiased momentum distribution. In the case of the 2012 data sample, this effect is not visible due to the larger number of candidates.
The parameters from the fits of Eq. (8.14) are listed in Tab. 8.10.

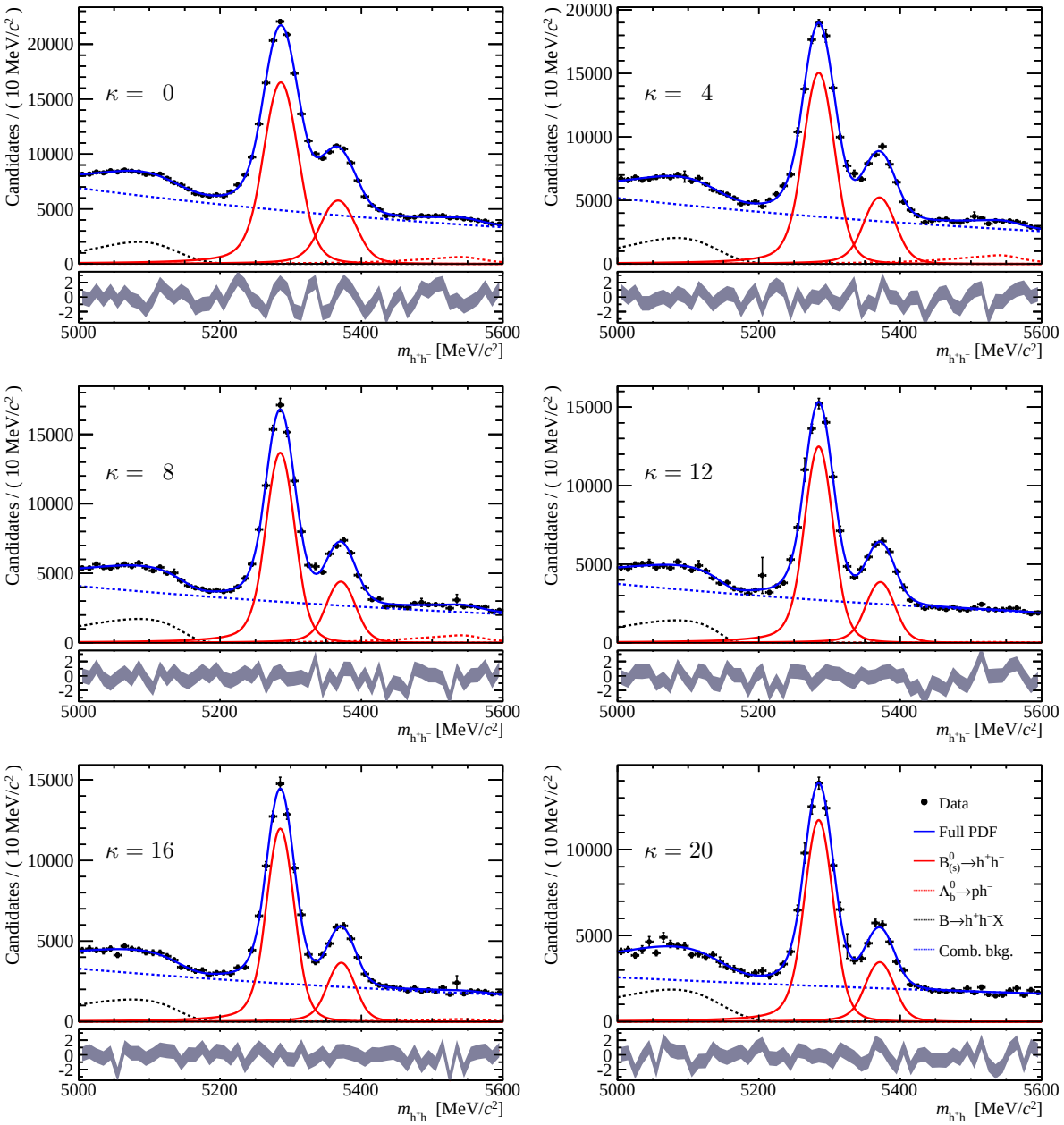


Figure 8.18: The invariant dihadron mass distribution of the decays $B^0_{(s)} \to h^+h'^-$ in 2011 data. The distribution is corrected for the PID efficiency and the plots correspond to different cut levels for the PID described by the parameter κ (*cf.* description in the text). The data is fit with two double Crystal Ball functions for the B^0 and B^0_s signal (red solid lines) and the sum of two Crystal Ball functions with common central value to describe the background from $\Lambda^0_b \to ph^-$ as described in Sec. 7.2 (red dotted line). Further there is a phenomenological function describing the contribution from partially reconstructed B decays (*cf.* Sec. 7.4.1; black dotted line) and an exponential function to describe the combinatorial background (blue dotted line). The grey band shows the pull.

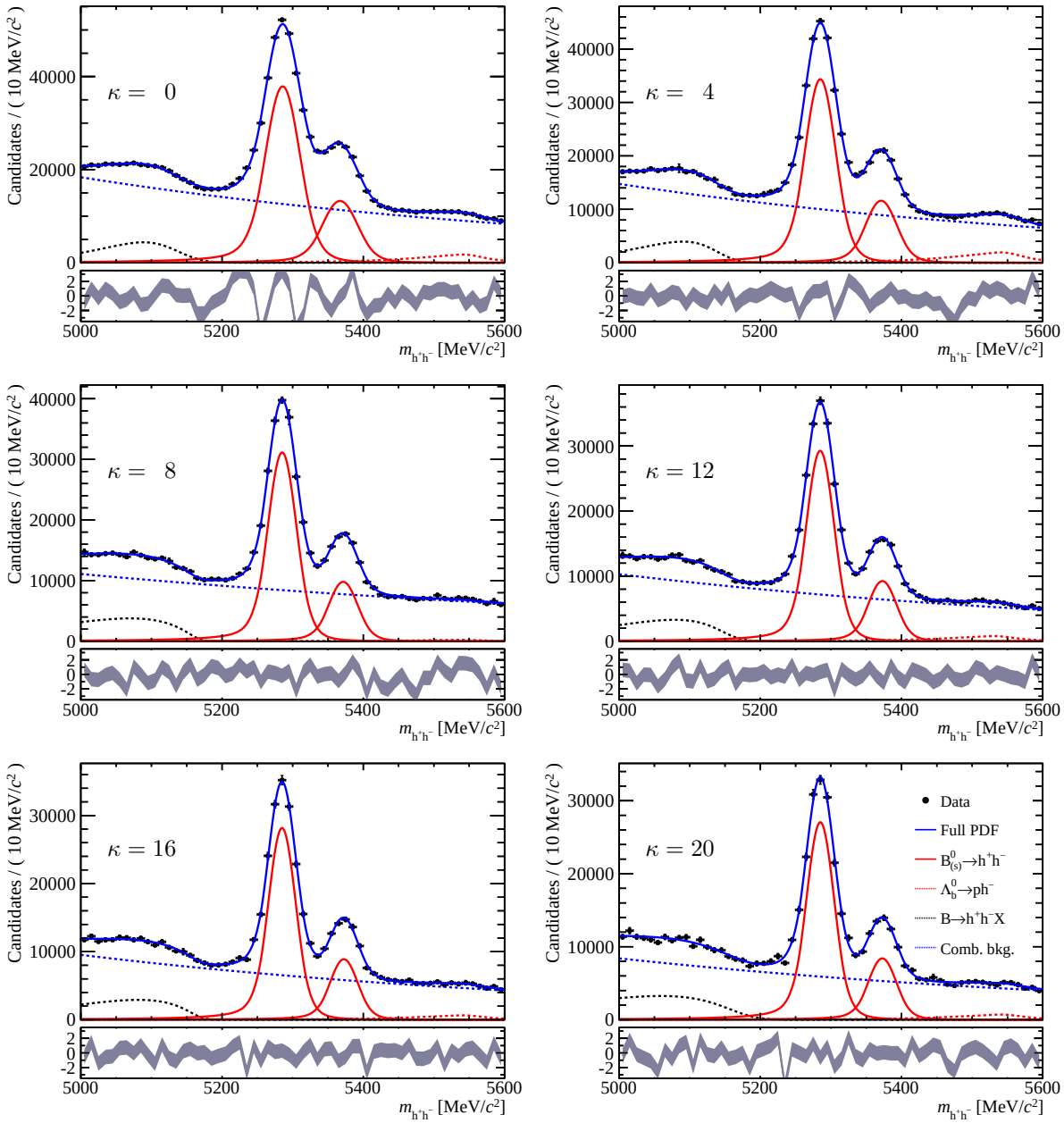


Figure 8.19: The invariant dihadron mass distribution of the decays $B^0_{(s)} \to h^+h'^-$ in 2012 data. The distribution is corrected for the PID efficiency and the plots correspond to different cut levels for the PID described by the parameter κ (*cf.* description in the text). The data is fit with two double Crystal Ball functions for the B^0 and B^0_s signal (red solid line) and the sum of two Crystal Ball functions with common central value to describe the background from $\Lambda^0_b \to ph^-$ as described in Sec. 7.2 (red dotted line). Further there is a phenomenological function describing the contribution from partially reconstructed B decays (*cf.* Sec. 7.4.1; black dotted line) and an exponential function to describe the combinatorial background (blue dotted line). The grey band shows the pull.

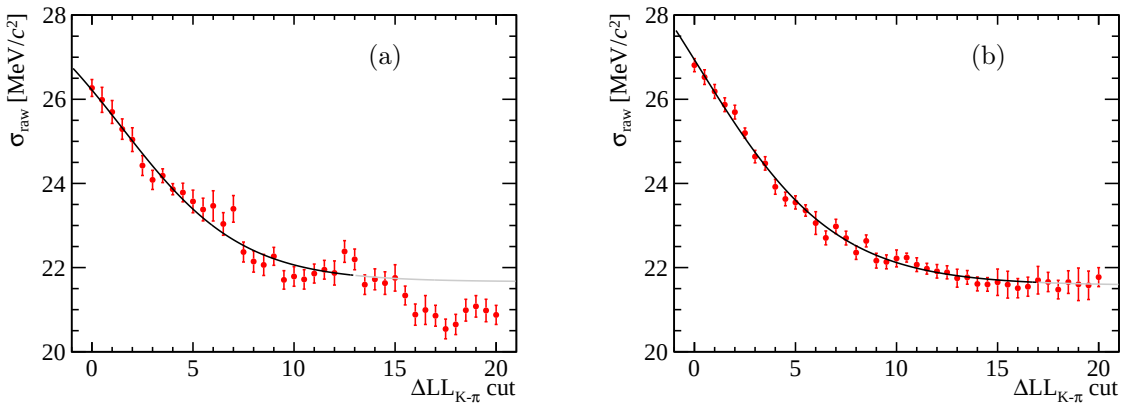


Figure 8.20: The width σ_{raw} of the double Crystal Ball function used in the description of the B^0 signal in the $m_{h^+h'^-}$ distribution as a function of the PID cut described by κ as explained in the text. σ_{raw} contains two components: one constant as a function of κ which is the invariant mass resolution and a second one varying as a function of κ due to the mis-identified final state particles. σ_{raw} is fit by the function described in Eq. (8.14). The black part shows the fit range and the grey part the continuation above the upper limit of the fit range.

Table 8.10: Fit parameters of $\sigma(\kappa)$ (described in Eq. (8.14)) for 2011 and 2012 data.

Data Set	$\boldsymbol{\sigma_{\text{mass-reso}}}$ **[MeV/c^2]**	$\boldsymbol{a_0}$ **[MeV/c^2]**	$\boldsymbol{\kappa_0}$	$\boldsymbol{r}$
2011 Data	21.66±0.18	9.4±5.3	0.31±0.08	1.05±0.34
2012 Data	21.57±0.10	10.1±4.3	0.31±0.05	0.94±0.21

As mentioned above the extracted values of $\sigma_{\text{mass-reso}}$ need to be corrected for the cut on the final state momenta of $5 < p < 200\,\text{GeV}/c$ as mentioned above. The correction factor, γ, defined as

$$\sigma = \gamma \cdot \sigma_{\text{mass-reso}}, \tag{8.15}$$

is obtained using two different approaches.
The first method uses $B_s^0 \to \mu^+\mu^-$ Monte Carlo. The invariant mass resolution is extracted as the width of the Gaussian kernel of a ${}_d\mathcal{CB}$ function fit to invariant dimuon mass distribution with and without this momentum cuts applied as shown in Fig. 7.4. The factor γ is estimated as the ratio of the resolution without and with the momentum cut.
The second method uses the interpolation method described above without and with the momentum cut applied to the two final state muons from the onia decays.
The resulting values of γ obtained using the two methods are listed in Tab. 8.11. The quoted uncertainties are statistical. The applied correction factor is the weighted average of the factors from the two methods, listed in the table. The corrected value for the invariant mass resolution

Table 8.11: Momentum range correction factor γ extracted from $B_s^0 \to \mu^+\mu^-$ MC and from the interpolation method. The total number is the weighted average.

Data Set	$\boldsymbol{B_s^0 \to \mu^+\mu^-}$ **MC**[a]	**Interpolation**	**Total**
2011 Data	1.030±0.006	1.032±0.014	1.030±0.005
2012 Data	1.030±0.006	1.032±0.012	1.030±0.005

[a] The same systematic uncertainty for both data sets is assigned as it is extracted from MC.

for B^0 is $22.30 \pm 0.19\,\text{MeV}/c^2$ for the 2011 data sample and $22.22 \pm 0.10\,\text{MeV}/c^2$ for the 2012 data sample. The uncertainties are statistical.
For B_s^0 these numbers have to be multiplied by the corresponding value of β_{reso} (*i.e.* the fixed ratio between the width of the B^0 and the B_s^0 signal in the fit) which gives an invariant mass resolution of $22.71 \pm 0.20\,\text{MeV}/c^2$ for the 2011 and $22.62 \pm 0.10\,\text{MeV}/c^2$ for the 2012 data sample, including the statistical uncertainty on β_{reso}.

Systematic Uncertainties

Six sources of systematic uncertainties have been considered: the uncertainty from the correction factor γ, the form of the fit function $\sigma_{\text{raw}}(\kappa)$, the effect of the fit range of $\sigma_{\text{raw}}(\kappa)$, the binning scheme for the determination of the PID efficiency, the uncertainty on the PID efficiency and the choice of the fit model to describe the invariant dihadron mass distributions.

1. **Correction factor $\boldsymbol{\gamma}$** The uncertainty of the factor γ displayed in Tab. 8.11 is propagated using standard error propagation.

2. **Fit function $\boldsymbol{\sigma(\kappa)}$** The uncertainty due to the choice in Eq. (8.14) is estimated by redoing the fit with an extended function including a third term proportional to $(1+r\cdot\exp(\kappa/\kappa_0))^{-2}$ to describe a possible contribution from candidates where both final state particles are misidentified. The pre-factor of the third term is $\mathcal{O}(10^{-2})\,\text{MeV}/c^2$. Thus, the change in the final value for $\sigma_{\text{mass-reso}}$ is also negligibly small.

3. **Fit range of $\boldsymbol{\sigma_{\text{raw}}(\kappa)}$** The uncertainty due to the chosen fit range is estimated by changing the upper boundary of the range by ± 0.5 (*i.e.* 2011 data: $[0, 12.5/13.5]$; 2012 data: $[0, 16.5/17.5]$). The maximal difference in the final value of σ using these modified boundaries is assigned as uncertainty. As expected, the uncertainty is larger in the 2011 data sample as the function has a larger absolute derivative at the upper boundary of the range.

4. **Binning scheme** The uncertainty from the binning scheme to determine the PID efficiency is estimated using the same modified binning scheme as described on page 122. The values for the mass resolution parameter are determined with this modified scheme and the differences to the nominal results are taken as systematic uncertainty.

5. **Uncertainty on the PID efficiency** The uncertainty due to the PID efficiency is estimated in a similar manner as for the BDT calibration. The difference in the $\Delta\text{LL}_{K-\pi}$ efficiency $\Delta\varepsilon(\kappa)$ estimated from the evaluation of $B^{\pm} \to J/\psi K^{\pm}$ candidates (*cf.* Fig. 7.10) is added to the $\Delta\text{LL}_{K-\pi}$ efficiency such that the weighting factor changes to

$$w = \Big([\varepsilon_{\text{kaon/pion}}(p_{h_i^+}, \eta_{h_i^+}, \text{n}_{\text{Tracks}}) + \Delta\varepsilon(\kappa)] \cdot [\varepsilon_{\text{kaon/pion}}(p_{h_i'^-}, \eta_{h_i'^-}, \text{n}_{\text{Tracks}}) + \Delta\varepsilon(\kappa)] \Big)^{-1}. \tag{8.16}$$

   The systematic uncertainty is defined as the difference of the final mass resolution parameter between the value obtained using this modified weighting factor and the nominal result.

6. **Fit function $\boldsymbol{m_{h^+h'^-}}$** The uncertainty due to the choice of the functional form used to describe the dihadron distribution is estimated by modifying it separately in two points. On the one hand the component for the combinatorial background is changed from a single exponential distribution to a sum of two exponential distributions. On the other hand the tail parameters for the B^0 and B_s^0 signal are not fixed to the values determined in MC, but are allowed to float in the range $\alpha_l \in [1, 5]$, $\alpha_r \in [-5, -1]$ and $n_{l,r} \in [4, 20]$.
   For both modifications, the mass resolution parameter is determined and the differences to the nominal result are added in quadrature to obtain the systematic uncertainty.

All systematic uncertainties are listed in Tab. 8.12. The total systematic uncertainty for each data sample is obtained by adding the individual uncertainties in quadrature.
The final estimates of the mass resolutions for $B^0_{(s)} \to \mu^+\mu^-$, obtained by studying the decays $B^0_{(s)} \to h^+h'^-$, are

$$\begin{aligned} \sigma_{B^0} &= 22.30 \pm 0.19(\text{stat.}) \pm 0.80(\text{syst.})\,\text{MeV}/c^2 \\ \sigma_{B_s^0} &= 22.71 \pm 0.19(\text{stat.}) \pm 0.82(\text{syst.})\,\text{MeV}/c^2 \end{aligned}$$

for 2011 data and

$$\begin{aligned} \sigma_{B^0} &= 22.22 \pm 0.10(\text{stat.}) \pm 0.71(\text{syst.})\,\text{MeV}/c^2 \\ \sigma_{B_s^0} &= 22.62 \pm 0.11(\text{stat.}) \pm 0.72(\text{syst.})\,\text{MeV}/c^2 \end{aligned}$$

for 2012 data.

Table 8.12: Systematic uncertainties on the mass resolution estimated with the decays $B^0_{(s)} \to h^+h'^-$.

Data Set	Systematic uncertainties [MeV/c^2]				
	Corr. fac. γ	Fit function $\sigma_{\mathrm{raw}}(\kappa)$	Fit range $\sigma(\kappa)$		
2011 Data	0.11	0.02	0.25		
2012 Data	0.11	0.03	0.02		
Data Set	**Systematic uncertainties [MeV/c^2]**				
	Binning scheme $\Delta LL_{K-\pi}$	**PID uncertainty**	**Fit function $m_{h^+h'^-}$**	**Total**	
2011 Data	0.45	0.50	0.34	0.80	
2012 Data	0.39	0.52	0.26	0.71	

8.2.3 Combination

The values obtained with the two methods (interpolation and $B^0_{(s)} \to h^+h'^-$) are in agreement within the uncertainties. They are combined by taking the weighted average, where the weighting factors are the inverse of the statistical and systematic uncertainties summed in quadrature, as there is no correlation of the systematic uncertainties between the two methods.
The combined values are

$$\sigma_{B^0} = 22.91 \pm 0.14(\text{stat.}) \pm 0.36(\text{syst.})\,\text{MeV}/c^2$$
$$\sigma_{B^0_s} = 23.28 \pm 0.15(\text{stat.}) \pm 0.41(\text{syst.})\,\text{MeV}/c^2$$

for 2011 data and

$$\sigma_{B^0} = 22.92 \pm 0.08(\text{stat.}) \pm 0.37(\text{syst.})\,\text{MeV}/c^2$$
$$\sigma_{B^0_s} = 23.32 \pm 0.09(\text{stat.}) \pm 0.38(\text{syst.})\,\text{MeV}/c^2$$

for 2012 data.
As the results are in good agreement between the 2011 and 2012 data samples, a combined value is calculated as the weighted average. Here, the statistical uncertainties are used as weights as the systematic uncertainties between the data samples are correlated. The combined systematic uncertainty is defined as the weighted average of the systematic uncertainties of the two data samples.
The obtained values are

$$\sigma_{B^0} = 22.92 \pm 0.07(\text{stat.}) \pm 0.37(\text{syst.})\,\text{MeV}/c^2$$
$$\sigma_{B^0_s} = 23.31 \pm 0.08(\text{stat.}) \pm 0.39(\text{syst.})\,\text{MeV}/c^2.$$

8.3 Tail Parameters

The tail parameters α and n of the single Crystal Ball function – used to describe the $B^0_{(s)} \to \mu^+\mu^-$ signal distribution – are estimated from $B^0_s \to \mu^+\mu^-$ MC as effects from final state radiation

(FSR) and bremsstrahlung are well-described in simulation. This determination is discussed in the second subsection, while an alternative method, also based on $B_s^0 \to \mu^+\mu^-$ MC, to estimate the tail parameters $\alpha_{l,r}$, $n_{l,r}$ and f of a ${}_d\mathcal{CB}$ function is discussed in the third subsection.

8.3.1 Phenomenology of the Tail Parameters

The non-Gaussian resolution effects described by the left-hand tail in the Crystal Ball distributions are mainly due to FSR, bremsstrahlung and other mechanisms of energy loss in the detector. The exponent n of this tail can be interpreted as the exponent describing the energy spectrum of final state radiation and bremsstrahlung, *i.e.*

$$\frac{\mathrm{d}N_\gamma}{\mathrm{d}E_\gamma} \propto E_\gamma^{-n}. \tag{8.17}$$

Theoretical studies predict that $n \approx 1$ [51].
There is a strong correlation between the tail parameters, especially the parameter α, describing the transition point between the Gaussian kernel and the exponential tails, and the width, σ of the Gaussian kernel. Thus the mass resolution of the sample used to estimate the tail parameters needs to be adjusted to that in the data.

The right-hand tail of the double Crystal Ball distributions describes non-Gaussian resolution effects caused by the detector and the reconstruction. Typically, these non-Gaussian effects are closer to the Gaussian limit (*i.e.* smaller deviation from a Gaussian resolution) than those in the left-hand tail, *i.e.* $n_r > n_l$ as the exponential tail transforms into Gaussian tails in the limit $n \to \infty$.

8.3.2 Tail Parameters for a Single Crystal Ball Function

Invariant Mass Distribution

The method applied to determine the tail parameters for the single Crystal Ball distribution uses the true four-vector momenta of the final state muons in the $B_s^0 \to \mu^+\mu^-$ MC sample. These four-vectors include the effect of FSR and bremsstrahlung and are used to calculate the true invariant dimuon mass distribution. This distribution is then convolved with a Gaussian distribution with a width σ_{smear} and the convolved distribution is fit with a single Crystal Ball to extract the width parameter σ_{fit} of the resulting distribution. Figure 8.21 shows as an example the fit distribution for $\sigma_{\mathrm{smear}} = 15\,\mathrm{MeV}/c^2$ and $25\,\mathrm{MeV}/c^2$ and Fig. 8.22 shows the width σ_{fit} as a function of σ_{smear}. There is a linear dependance and a regression of $\sigma_{\mathrm{fit}}(\sigma_{\mathrm{smear}}) = a_0 + a_1 \cdot \sigma_{\mathrm{smear}}$ yields $a_0 = 0.974 \pm 0.007\,\mathrm{MeV}/c^2$ and $a_1 = 0.9955 \pm 0.0006$. As expected a_1 is smaller than one as the true invariant dimuon mass distribution incorporates the intrinsic width of the signal as well as the width due to energy loss effects in the detector.
The values of σ_{smear} obtained by the regression for B^0 and B_s^0 in the two data sets are listed in Tab. 8.13. The quoted uncertainties are due to uncertainties on the previously estimated values of σ as well as on the regression itself.

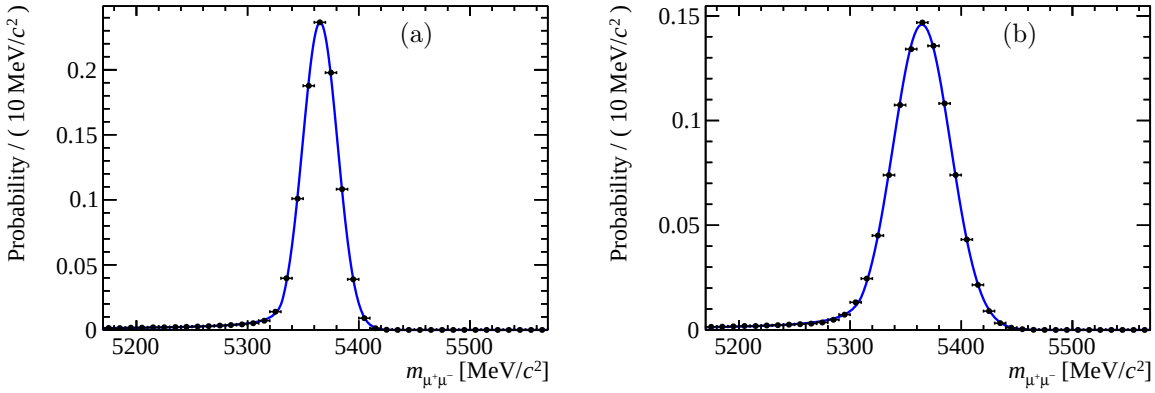


Figure 8.21: The invariant dimuon mass distribution based on the true four-vector momentum of the muons in $B_s^0 \to \mu^+\mu^-$ MC. The generated distributions are convolved with a Gaussian distribution with a width of (a) 15 MeV/c^2 and (b) 25 MeV/c^2. The blue line shows the fitted Crystal Ball function.

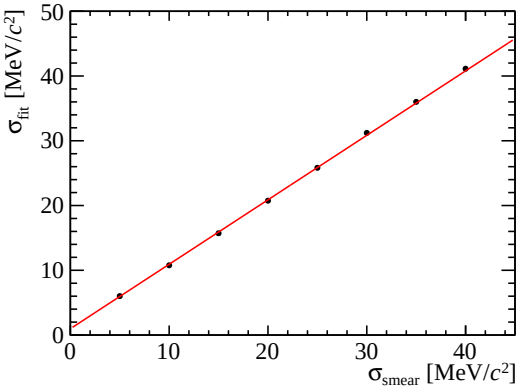


Figure 8.22: The resulting invariant dimuon mass resolution σ_{fit} as a function of the width of the Gaussian σ_{smear} used to smear the invariant dimuon mass distribution based on the true four-vector momentum of the muons in $B_s^0 \to \mu^+\mu^-$ MC. The points are fit with a linear function.

Table 8.13: Values of σ_{smear} to reproduce the invariant mass resolution in data.

Data Set	$\boldsymbol{\sigma}_{\text{smear}}$ **[MeV/c^2]**	
	$\boldsymbol{B^0}$	$\boldsymbol{B_s^0}$
2011	22.04 ± 0.50	22.41 ± 0.54
2012	22.05 ± 0.45	22.45 ± 0.47

Extraction of the Parameters

For each value of σ_{smear}, the invariant dimuon mass distribution based on the true four-vector momentum of the muons convolved with a Gaussian with mean zero and width σ_{smear} is fit with a single Crystal Ball function (*cf.* Fig. 8.23). The extracted parameters α and n are listed in Tab. 8.14 and Fig. 8.24 shows the correlation matrix for all four parameters of the Crystal Ball distribution.

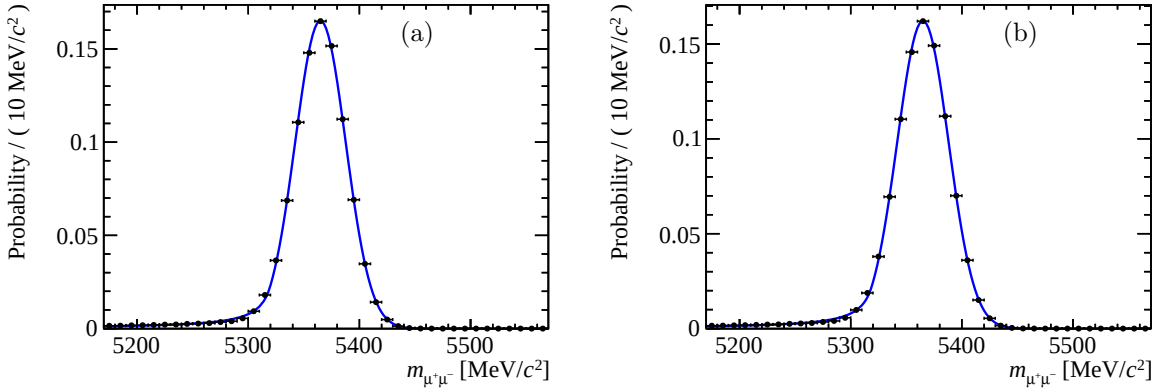


Figure 8.23: The invariant dimuon mass distributions based on the true four-vector momentum of the muons in $B_s^0 \to \mu^+\mu^-$ MC and convolved with a Gaussian to reproduce the mass resolution shown as an example for (a) $B^0 \to \mu^+\mu^-$ in 2011 data and (b) $B_s^0 \to \mu^+\mu^-$ in 2012 data. The blue line shows the fit of the Crystal Ball function.

Systematic Uncertainties

The systematic uncertainties of the parameters α and n are estimated as the standard deviations of the parameter distributions (*cf.* Fig. 8.25) given by the uncertainty on σ_{smear}. These distributions are generated by Toy Monte Carlo, where the extraction of the parameters is repeated 1000 times while σ_{smear} is allowed to float according to its uncertainty treated as a Gaussian constraint.
The final values obtained for α and n are listed in Tab. 8.14. The parameters between 2011 and 2012 data are in good agreement and their weighted average based on the statistical uncertainty is taken. The systematic uncertainties are correlated between the two data samples and the arithmetic average is taken for the combination. The obtained values are for α

$$\begin{aligned} \alpha_{B^0} &= 2.138 \pm 0.005(\text{stat.}) \pm 0.012(\text{syst.}) \\ \alpha_{B_s^0} &= 2.128 \pm 0.006(\text{stat.}) \pm 0.011(\text{syst.}) \end{aligned}$$

and for n

$$\begin{aligned} n_{B^0} &= 0.925 \pm 0.010(\text{stat.}) \pm 0.020(\text{syst.}) \\ n_{B_s^0} &= 0.936 \pm 0.011(\text{stat.}) \pm 0.020(\text{syst.}). \end{aligned}$$

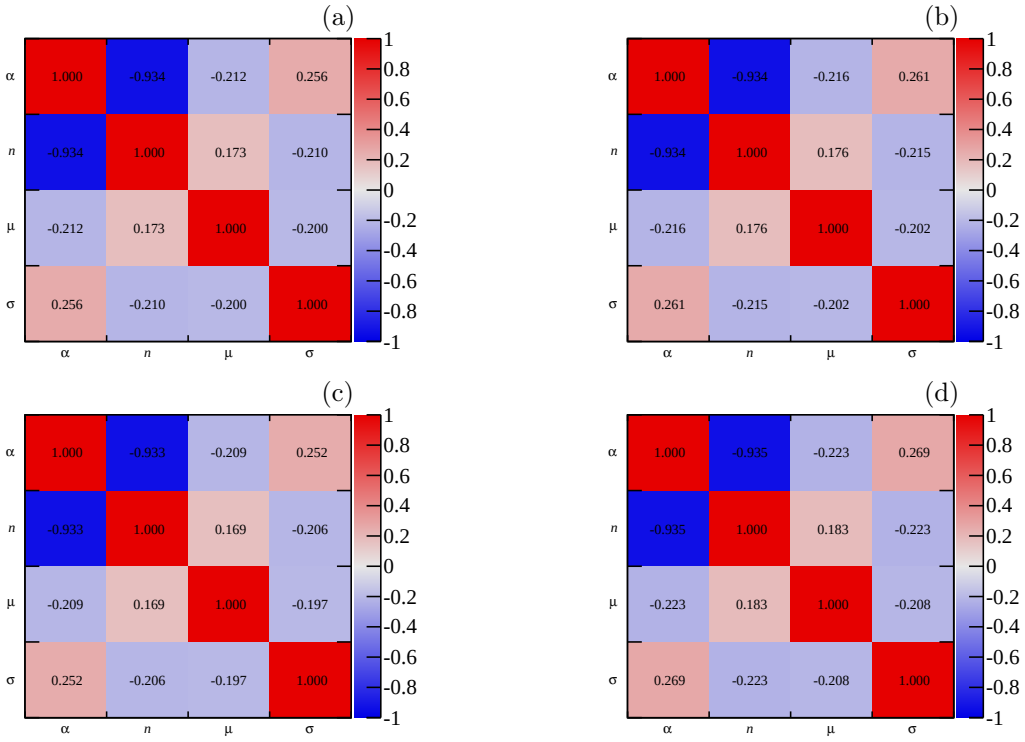


Figure 8.24: Correlation matrices of the four parameters in the single Crystal Ball function for (a) $B^0 \to \mu^+\mu^-$ and (b) $B_s^0 \to \mu^+\mu^-$ in 2011 data as well as (c) $B^0 \to \mu^+\mu^-$ and (d) $B_s^0 \to \mu^+\mu^-$ in 2012 data. As expected there is a large anti-correlation between α and n. Furthermore, α and σ are positively correlated as μ and σ are negatively correlated due to the radiative tail towards lower mass values.

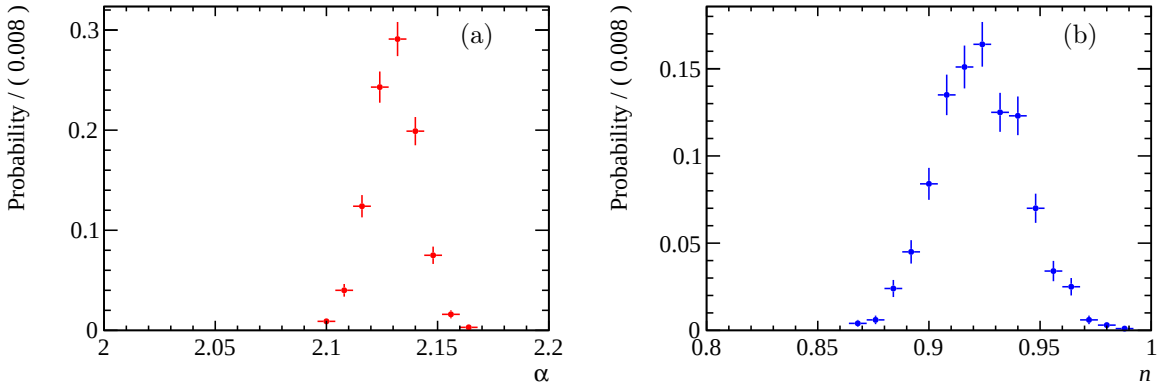


Figure 8.25: The distribution of (a) α and (b) n for B_s^0 in 2012 data based on the uncertainty of σ_{smear} due to the uncertainty on the mass resolution. The standard deviation of the distributions is used as systematic uncertainty on the values of α and n. All distributions for α and n for the 2011 data as well as B^0 in 2012 are shown in Appendix B.

Table 8.14: Values of α and n for B^0 and B_s^0 in 2011 and 2012 data. The first uncertainty is statistical, the second one systematic.

Data Set	$\boldsymbol{\alpha}$	$\boldsymbol{n}$
2011		
B^0	$2.139 \pm 0.007 \pm 0.012$	$0.924 \pm 0.015 \pm 0.020$
B_s^0	$2.124 \pm 0.008 \pm 0.011$	$0.926 \pm 0.014 \pm 0.020$
2012		
B^0	$2.137 \pm 0.007 \pm 0.011$	$0.926 \pm 0.014 \pm 0.020$
B_s^0	$2.132 \pm 0.008 \pm 0.011$	$0.935 \pm 0.016 \pm 0.020$

8.3.3 Tail Parameters for a Double Crystal Ball Function

The right-hand tail of the double Crystal Ball distribution describes non-Gaussian detector resolution effects. Such effects are not included in the invariant dimuon mass distribution based on the true four-vector momentum of the muons, but only appear in the reconstructed invariant dimuon mass distribution.
As discussed above, the invariant dimuon mass resolution observed in MC is not in agreement with that in data and needs therefore to be adapted to the latter. To disentangle the component from bremsstrahlung and other mechanisms of energy loss from the component due to the detector resolution, the invariant dimuon mass distribution $m_{\mu^+\mu^-,\text{truth}}$ based on the true four-momentum of the muons and the distribution of the difference between the reconstructed and true dimuon masses $\Delta m = m_{\mu^+\mu^-,\text{rec}} - m_{\mu^+\mu^-,\text{truth}}$ (*cf.* Fig. 8.26) in simulation are studied.
The distribution of Δm is scaled by a factor β to reproduce the observed invariant mass resolution in data. The corresponding values of β are estimated by fitting a ${}_d\mathcal{CB}$ function to the distributions $m_{\mu^+\mu^-,\text{truth}} + \beta \cdot \Delta m$, where β is considered in the range from 0 to 2 in steps of 0.2. Figure 8.27 shows the fit distribution for $\beta = 0.2$ and $\beta = 1.6$. The obtained mass resolution parameter σ_{fit} as a function of β is shown in Fig. 8.22. The data points are fit with a linear function $\sigma(\beta) = b_0 + b_1 \cdot \beta$. The obtained values of the fit parameters are $b_0 = 1.361 \pm 0.008\,\text{MeV}/c^2$ and $b_1 = 21.004 \pm 0.017\,\text{MeV}/c^2$.
The values of β corresponding to the measured values of σ for B^0 and B_s^0 in the two data samples are listed in Tab. 8.15. The quoted uncertainties are due to uncertainties on σ as well as on the regression.

Table 8.15: Values of β to reproduce the invariant mass resolution in data.

Data Set	$\boldsymbol{\beta}$	
	$\boldsymbol{B^0}$	$\boldsymbol{B_s^0}$
2011	1.026 ± 0.024	1.044 ± 0.027
2012	1.027 ± 0.021	1.046 ± 0.022

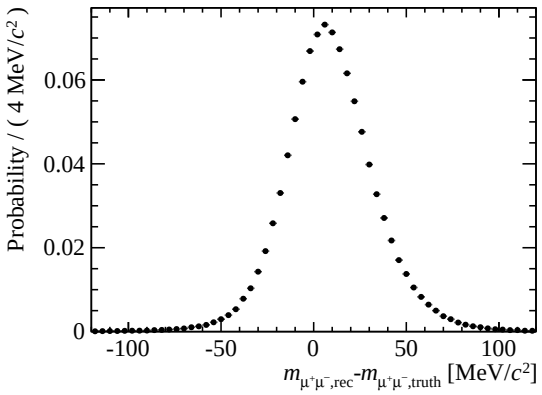


Figure 8.26: Distribution of the mass difference, Δm, between the reconstructed and the true invariant dimuon mass in $B_s^0 \to \mu^+\mu^-$ MC.

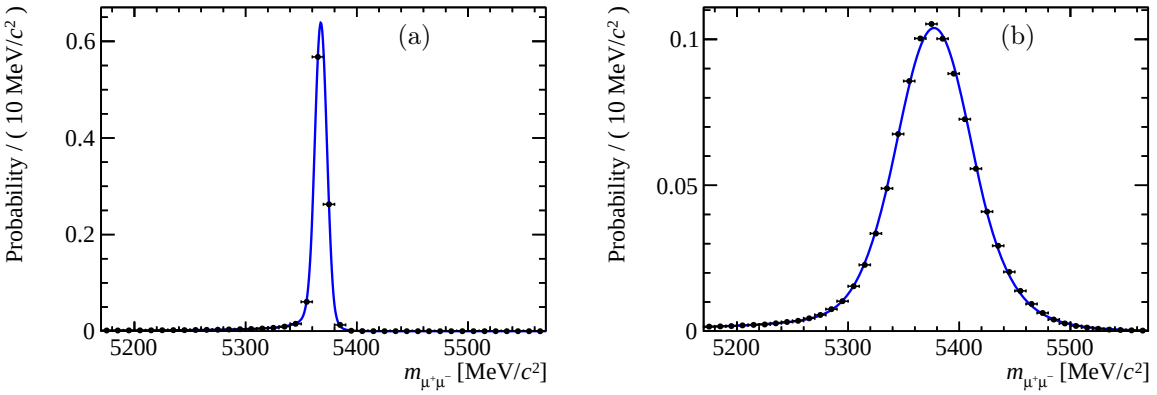


Figure 8.27: The invariant dimuon mass distribution based on the true four-vector momentum of the muons in $B_s^0 \to \mu^+\mu^-$ MC. The distribution is convolved with the distribution of the differences, Δm, between the reconstructed and true dimuon mass, scaled by a factor β which is (a) 0.2 and (b) 1.6. The blue line shows the fit ${}_d\mathcal{CB}$ function.

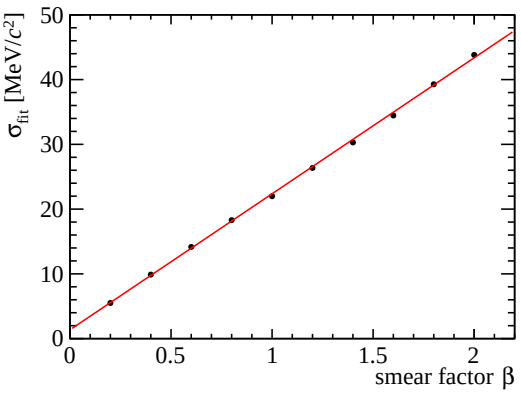


Figure 8.28: The resulting invariant dimuon mass resolution σ_{fit} as a function of the scale factor β on the distribution of the difference, Δm, between the reconstructed and true dimuon mass in $B_s^0 \to \mu^+\mu^-$ MC. The points are fit by a linear function.

Extraction of Parameters

The distributions $m_{\mu^+\mu^-,\text{truth}} + \beta \cdot \Delta m$ with the values of β determined above are fit with a ${}_d\mathcal{CB}$ function using a binned maximum-likelihood method (*cf.* Fig. 8.29) to extract the tail parameters $\alpha_{l,r}$, $n_{l,r}$ and f as well as the correlation matrix of the seven parameters. The parameters are listed in Tab. 8.16 and the correlation matrices are shown in Fig. 8.30.

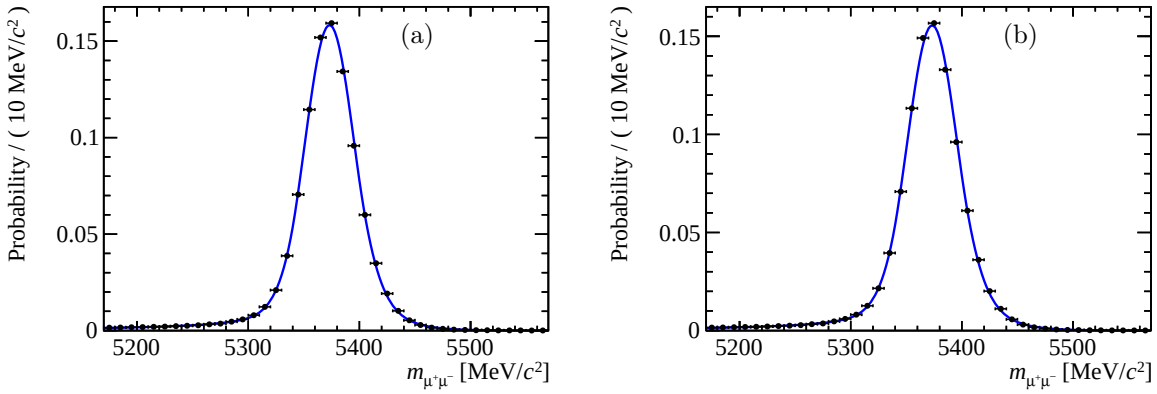


Figure 8.29: The invariant dimuon mass distributions based on the true four-vector momentum of the muons in $B_s^0 \to \mu^+\mu^-$ MC and the scaled difference between reconstructed and true dimuon mass to reproduce the mass resolution shown as an example for (a) $B^0 \to \mu^+\mu^-$ in 2011 data and (b) $B_s^0 \to \mu^+\mu^-$ in 2012 data. The blue line shows the fitted ${}_d\mathcal{CB}$ function.

Systematic Uncertainties

The systematic uncertainties of the parameters f, $\alpha_{l,r}$ and $n_{l,r}$ are estimated as the standard deviations of the parameter distributions (*cf.* Fig. 8.31) given the uncertainty on β. These distributions are generated by Toy Monte Carlo, where the extraction of the parameters is

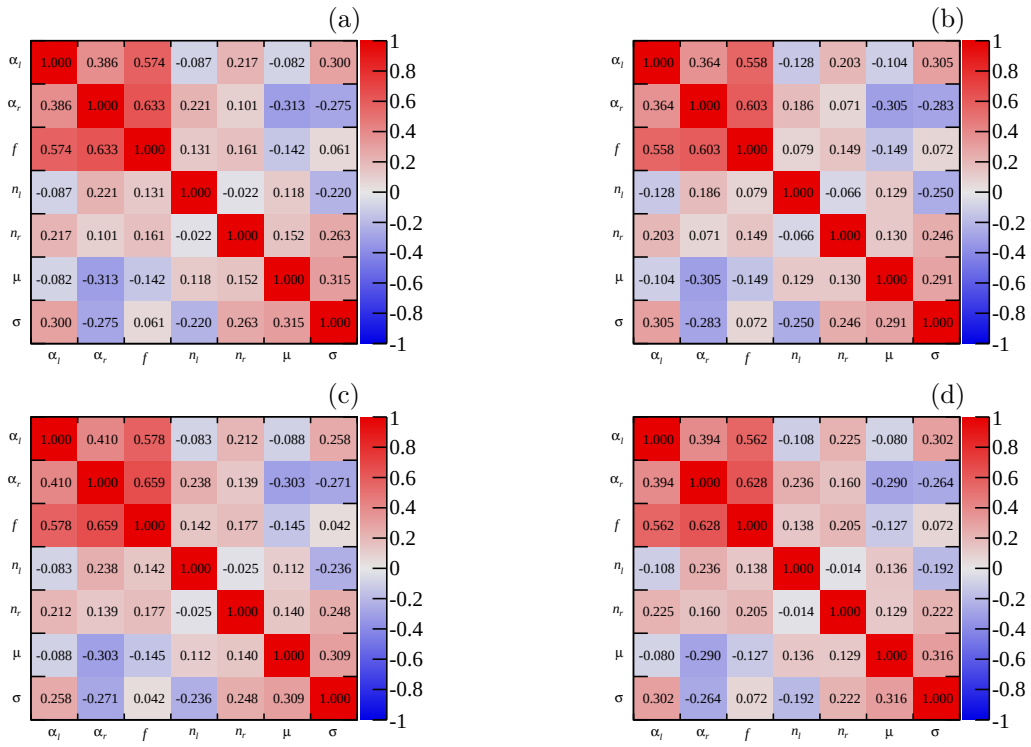


Figure 8.30: Correlation matrix between the seven parameters in the double Crystal Ball function for (a) $B^0 \to \mu^+\mu^-$ and (b) $B_s^0 \to \mu^+\mu^-$ in 2011 data as well as (c) $B^0 \to \mu^+\mu^-$ and (d) $B_s^0 \to \mu^+\mu^-$ in 2012 data. There are large correlations among the transition point parametes $\alpha_{l,r}$ and the parameter f. Further μ and σ are – in contrast to the single $\mathcal{CB}$ case (*cf.* Fig. 8.24) – positively correlated due to the additional tail in the distribution to the right-hand side.

Table 8.16: Values of $n_{l,r}$, $\alpha_{l,r}$ and f in 2011 and 2012 data. The first uncertainty is statistical, the second one systematic.

Data Set			
	$\boldsymbol{\alpha_l}$	$\boldsymbol{\alpha_r}$	
2011			
B^0	$1.354 \pm 0.012 \pm 0.016$	$-1.136 \pm 0.006 \pm 0.004$	
B_s^0	$1.352 \pm 0.007 \pm 0.016$	$-1.135 \pm 0.008 \pm 0.005$	
2012			
B^0	$1.355 \pm 0.009 \pm 0.014$	$-1.132 \pm 0.007 \pm 0.005$	
B_s^0	$1.349 \pm 0.008 \pm 0.017$	$-1.134 \pm 0.008 \pm 0.004$	
	$\boldsymbol{n_l}$	$\boldsymbol{n_r}$	
2011			
B^0	$1.362 \pm 0.023 \pm 0.018$	142.9 ± 0.4	± 4.7
B_s^0	$1.379 \pm 0.020 \pm 0.020$	143.6 ± 0.5	± 4.8
2012			
B^0	$1.359 \pm 0.023 \pm 0.016$	143.7 ± 0.9	± 4.7
B_s^0	$1.365 \pm 0.024 \pm 0.019$	143.0 ± 0.6	± 4.8
	$\boldsymbol{f}$		
2011			
B^0	$0.418 \pm 0.019 \pm 0.009$		
B_s^0	$0.415 \pm 0.021 \pm 0.008$		
2012			
B^0	$0.421 \pm 0.016 \pm 0.007$		
B_s^0	$0.415 \pm 0.017 \pm 0.007$		

repeated 1000 times while β is allowed to float according to its uncertainty treated as a Gaussian constraint. The final parameters are listed in Tab. 8.16.
As the parameters for 2011 data and 2012 data are in good agreement, the values of the 2011 and 2012 data samples are combined in the same manner as described above for the single Crystal Ball parameters. The combined values are

$$
\begin{aligned}
\alpha_{l,B^0} &= 1.355 \pm 0.007(\text{stat.}) \pm 0.015(\text{syst.}) \\
\alpha_{l,B_s^0} &= 1.350 \pm 0.005(\text{stat.}) \pm 0.016(\text{syst.}) \\
\alpha_{r,B^0} &= -1.134 \pm 0.005(\text{stat.}) \pm 0.005(\text{syst.}) \\
\alpha_{r,B_s^0} &= -1.134 \pm 0.006(\text{stat.}) \pm 0.005(\text{syst.}) \\
n_{l,B^0} &= 1.360 \pm 0.016(\text{stat.}) \pm 0.018(\text{syst.}) \\
n_{l,B_s^0} &= 1.374 \pm 0.015(\text{stat.}) \pm 0.018(\text{syst.}) \\
n_{r,B^0} &= 143.0 \pm 0.4 \ (\text{stat.}) \pm 4.7 \ (\text{syst.}) \\
n_{r,B_s^0} &= 143.2 \pm 0.4 \ (\text{stat.}) \pm 4.8 \ (\text{syst.}) \\
f_{B^0} &= 0.420 \pm 0.012(\text{stat.}) \pm 0.008(\text{syst.}) \\
f_{B_s^0} &= 0.417 \pm 0.013(\text{stat.}) \pm 0.007(\text{syst.}).
\end{aligned}
$$

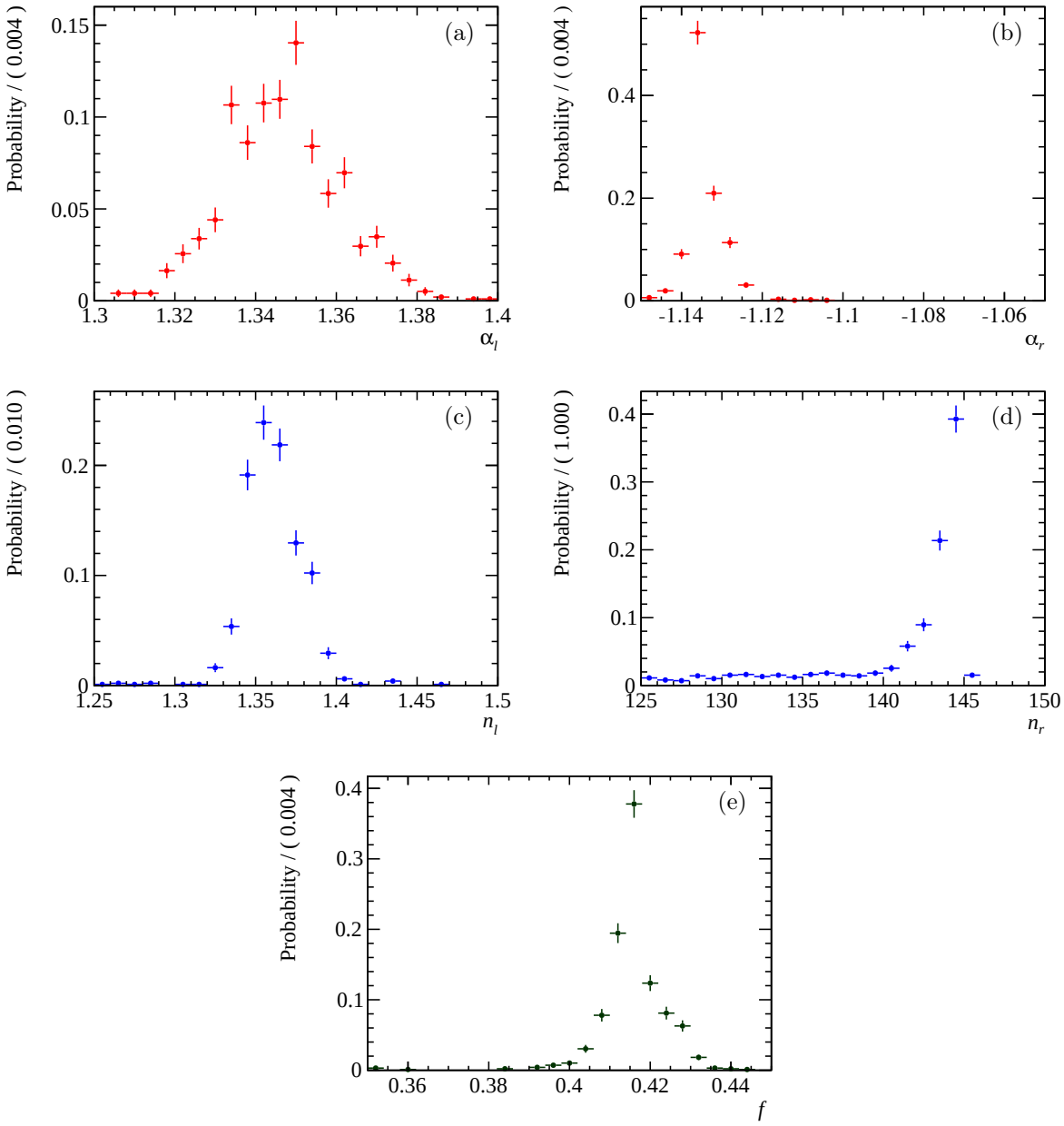


Figure 8.31: The distribution of (a) α_l, (b) α_r, (c) n_l, (d) n_r and (e) f for B_s^0 in 2012 data based on the uncertainty of β due to the uncertainty on the mass resolution. The standard deviation of the distributions is used as systematic uncertainty on the parameters. All distributions of the parameters for the 2011 data as well as B^0 in 2012 are shown in Appendix B.

9

Latest Results on $B^0_{(s)} \to \mu^+\mu^-$ and Their Interpretation

The three LHC experiments, LHCb, CMS and ATLAS, have published several searches and branching fraction measurements of the decays $B^0_{(s)} \to \mu^+\mu^-$ since the start of the LHC. Furthermore, the two TeVatron experiments CDF and D0 have presented updated results. This chapter presents in the first section the latest experimental results and in the second section their implications.

9.1 Experimental Results

As the first LHC experiment, LHCb published upper limits on $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ based on a data sample corresponding to the 37 pb^{-1} collected in 2010 [19]. These first LHCb limits were a factor 4 and 10 above the then-current best upper limits by CDF [3]. Improved LHCb measurements based on a sample corresponding to 0.37 fb^{-1} [21] and 1 fb^{-1} [22] resulted in the World's best upper limits. The limits at 95 % C.L. based on 1 fb^{-1} were a factor 1.3 for $B^0_s \to \mu^+\mu^-$ and 10 for $B^0 \to \mu^+\mu^-$ above the SM prediction.
The CMS experiment published its first upper limits on the two channels based on a data sample of 1 fb^{-1} [109] and updated it using the full 2011 data sample, corresponding to 5 fb^{-1} [112].
A first upper limit by ATLAS on $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ was based on about half of its 2011 data sample, corresponding to 2.4 fb^{-1} [13] and was later updated with the full 2011 data sample of 5 fb^{-1} [7].
In the meantime also the two TeVatron experiments, CDF and D0, presented updated results. Based on a data sample corresponding to 7 fb^{-1}, CDF published a first two sided confidence limit for $B^0_s \to \mu^+\mu^-$ whose central value was a factor six above the SM prediction [37]. The upper limits on $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ quoted in Ref. [37] were less stringent than the results published by LHCb [21] and CMS [109] almost at the same time. The CDF result was later updated using the full data sample collected in the TeVatron Run II, corresponding to

10 fb^{-1} [38], and also D0 published a result for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ using its full Run II data sample of 10 fb^{-1} [41].

The first evidence of the decay $B_s^0 \to \mu^+\mu^-$ was seen by LHCb based on a combined analysis of the full 2011 data sample and about half of the 2012 data sample, corresponding to $1.0 + 1.1$ fb^{-1} [24]. The observed statistical significance of the signal was 3.5σ. This result was confirmed using the full 2011 and 2012 data samples ($1 + 2$ fb^{-1}) with an observed statistical significance of 4.0σ for $B_s^0 \to \mu^+\mu^-$ while it was 2.0σ for $B^0 \to \mu^+\mu^-$. The measured branching fractions were $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.9^{+1.2}_{-1.2}) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.7^{+2.5}_{-2.1}) \times 10^{-10}$ where the uncertainties are statistical and systematic combined. The upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ was 7.4×10^{-10} at 95 % C.L. [25].
The corresponding results from CMS, also based on the full 2011 and 2012 data samples ($5 + 20$ fb^{-1}), gave significances of 4.3σ for $B_s^0 \to \mu^+\mu^-$ and 2.0σ for $B^0 \to \mu^+\mu^-$. The measured branching fractions were $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}) \times 10^{-10}$, where the uncertainties are again statistical and systematic combined. The upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ was 1.1×10^{-9} at 95 % C.L. [114].

The preliminary combination of the latest measurements by LHCb and CMS [6] led to measured branching fractions of

$$\begin{aligned} \mathcal{B}(B_s^0 \to \mu^+\mu^-) &= (2.9^{+0.7}_{-0.7}) \times 10^{-9} \\ \mathcal{B}(B_s^0 \to \mu^+\mu^-) &= (3.6^{+1.6}_{-1.4}) \times 10^{-10}, \end{aligned}$$

where the quoted uncertainty is the statistical and systematic uncertainties added in quadrature. Figure 9.1 shows the latest results of the TeVatron and LHC experiments as well as the preliminary combination of the latest LHCb and the CMS results and the SM predictions.

9.2 Implications

The LHCb and CMS measurements of the branching fraction of the decay $B_s^0 \to \mu^+\mu^-$ are in good agreement with the SM predictions and put severe constraints on many models describing physics beyond the SM.
The constraints on CMSSM are particular strong at large values of $\tan\beta$. For $\tan\beta = 50$ and $A_0 = -2m_0$, the measurements disallow a mass of the fermions at the GUT scale up to $m_{1/2} = 1250\,\mathrm{GeV}/c^2$ [64, 194]. These constraints are much stronger than those from direct searches by CMS and ATLAS. On the other hand, the limits on $m_{1/2}$ cannot compete with direct searches in the region of $\tan\beta$ below 30. In this case, also other indirect searches such as in $b \to s\gamma$ or in $B^0 \to K^{*0}\mu^+\mu^-$ have stronger constraints than $B_s^0 \to \mu^+\mu^-$ [194].
Similar constraints in terms of $m_{1/2}$ as in the CMSSM are set by the $\mathcal{B}(B_{(s)}^0 \to \mu^+\mu^-)$ measurements in the NUHM1 model with $\tan\beta = 30$, $m_A = 1$ TeV/c^2, $A_0 = 2.5m_0$ and $\mu > 0$ and in minimal supergravity (mSUGRA) models with $A_0/m_0 > 2$ and $\mu > 0$ [94].
A detailed description of the impact of the $\mathcal{B}(B_{(s)}^0 \to \mu^+\mu^-)$ measurements, in combination with complementary measurements from LHC (*e.g.* missing transverse energy measurements, Higgs boson mass), but also with astrophysical measurements (*e.g.* WIMP-searches, dark matter density) on the CMSSM and NUHM1 parameter space can be found in Ref. [93].

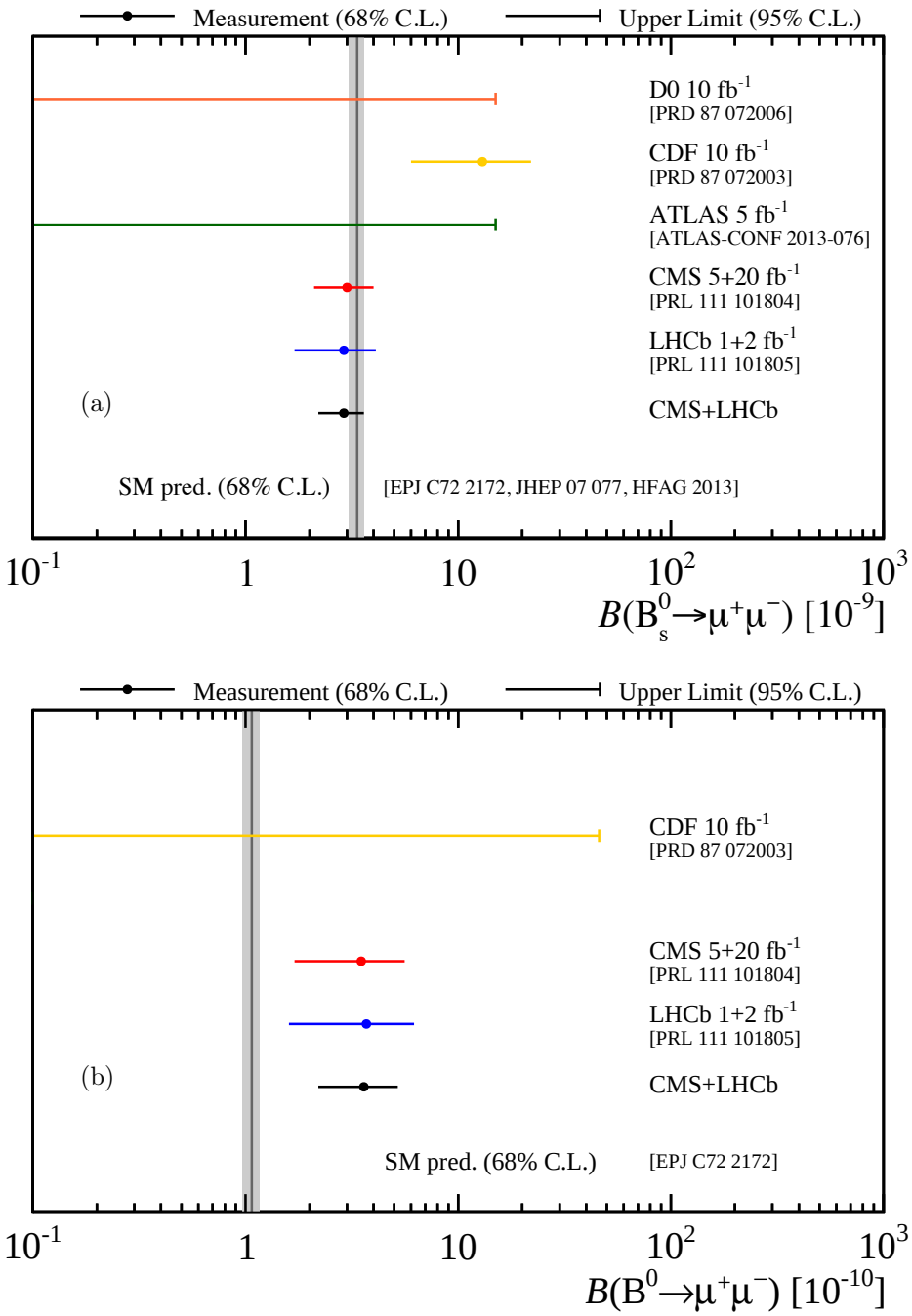


Figure 9.1: Latest results on (a) $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and (b) $\mathcal{B}(B^0 \to \mu^+\mu^-)$ of the TeVatron and LHC experiments as well as the combination of the latest LHCb and the CMS results and the SM predictions.

A particular point that has raised interest, is the measured enhancement of the branching fraction $\mathcal{B}(B^0 \to \mu^+\mu^-)$. Already the analysis that led to the first evidence of $B_s^0 \to \mu^+\mu^-$ [24] showed a higher $B^0 \to \mu^+\mu^-$ branching fraction than expected. The measurements using the full LHC Run I data samples by LHCb [25] and CMS [114] show both an enhancement of the branching fraction by a factor of about 3. Both measurements deviate from the SM prediction by more than 1σ. Such an enhancement could be explained for example by a fourth quark generation with $V_{t'd} > V_{t's}$ as $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ has been measured to be SM-like. But such a scenario would not be compatible with the observation of an SM-like Higgs boson at a mass of about $125\,\mathrm{GeV}/c^2$ [171].

Further, the measurement of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ also puts constraints on the tcZ coupling. With $\mathcal{B}(t \to Zc) < 6.3 \times 10^{-5}$, these constraints are more stringent than the individual 5σ discovery potential of CMS or ATLAS with data samples corresponding to 10 fb^{-1} [161].

IV

Z Production in Proton-Lead Collisions

Mehr als das Gold hat das Blei die Welt verändert.
Und mehr als das Blei in der Flinte das im Setzkasten.

Georg Christoph Lichtenberg (German physicist and author, 1742-1799)

10

Theory of Z Production in Proton-Nucleus Collisions

Studies of proton-nucleus collisions – done at the LHC colliding a proton beam and a $^{208}_{82}$Pb beam – are a crucial component of the physics program with high-energy nuclear beams. In particular, properties measured in proton-nucleus collisions serve as reference to interpret and understand nucleus-nucleus collisions. Another important feature of proton-nucleus collisions is their potential to reveal the partonic structure of matter described by *nuclear parton distribution functions* (nPDF), especially at low parton fractional momenta (small Bjorken-x) [213]. A measurement that might serve as input for the determination of the partonic structure of nuclei is the production of Z bosons in proton-nucleus collisions. This measurement, performed with Z bosons decaying into opposite-charged muons at LHCb, is described in the next chapter.
The first section of this chapter describes the proton-lead collisions performed at the LHC. The second section is devoted to a short introduction to nPDFs while the last section describes the production of Z bosons in hadron colliders including predictions for this process.

10.1 Proton-Lead Collisions at the LHC measured by the LHCb Experiment

A data set of proton-lead collisions has been collected by the LHC experiments at the beginning of the year 2013. Due to the asymmetric detectors of LHC – ALICE and LHCb – data has been taken with the lead beam propagating clock- and counter-clockwise. The configuration where the proton beam propagates in the positive z-direction in LHCb, *i.e.* the proton flies from the region of the interaction point to the region of the muon stations, is referred to as pA while the other configuration is called Ap[1].
The proton beam had a nominal energy of $E_{\text{proton}} = 4\,\text{TeV}$ while the nuclear beam, consisting

[1] The first configuration is also named 'forward' (*i.e.* in the direction of the proton beam) and the second one 'backward' (*i.e.* against the direction of the proton beam). The rapidity is always measured with respect to the proton beam. The pseudorapidity is always measured in the lab frame of the LHCb detector.

of lead nuclei with $N_Z = 82$ protons and $N_A = 208$ nucleons in total, was operated at an energy per nucleon of $E_{\text{nucleon}} = N_Z/N_A \cdot E_{\text{proton}} \approx 1.58\,\text{TeV}$, leading to a center-of-mass energy per proton-nucleon pair of $\sqrt{s_{NN}} = 5\,\text{TeV}$ [2].
There is a Lorentz boost between the lab and the centre-of-mass frame due to the asymmetric beam energies. The difference in rapidity ($y = 1/2 \cdot \ln[(E + p_z)/(E - p_z)]$) between the lab frame, y_{Lab}, and the centre-of-mass frame, y, is $\Delta y = y_{\text{Lab}} - y = 1/2 \log(A/Z) = 0.47$, where the rapidity is always measured with respect to the proton beam.
During the proton-lead data taking period, LHCb has collected about 2 nb^{-1} of integrated luminosity whereof 1.4 nb^{-1} were collected in the pA configuration and the other 0.6 nb^{-1} in the Ap configuration. Figure 10.1 shows the integrated luminosity as a function of time for the beginning of the year 2013.

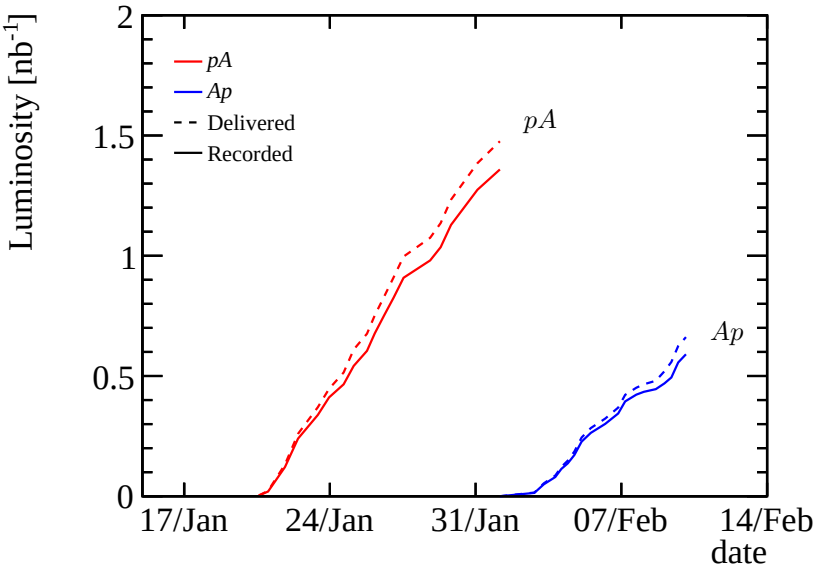


Figure 10.1: Delivered and recorded integrated luminosity in pA and Ap runs at LHCb as a function of time in the year 2013.

10.2 Nuclear Parton Distribution Functions

10.2.1 General Aspects of Nuclear Parton Distribution Functions

The general ansatz to compute inclusive cross sections of hard processes in hadron collisions, where the interaction scale $Q \gg \Lambda_{\text{QCD}}$ is large, is the factorisation theorem

$$\mathrm{d}\sigma_{AB\to C+X} = \sum_{a,b} f_a^A(x, Q^2) \cdot f_b^B(x, Q^2) \cdot \mathrm{d}\hat{\sigma}_{ab\to C+X}\,, \tag{10.1}$$

[2] The measured proton beam energy for the 2012 configuration is $E_p = 3988 \pm 26$ GeV leading to an energy of the lead beam per nucleon of $E_N = 1572 \pm 10$ GeV. This gives a centre-of-mass energy per proton-nucleon pair of $\sqrt{s_{NN}} = 5008 \pm 33$ GeV, which is approximated to 5 TeV [232].

where A and B are the colliding particles, $\mathrm{d}\hat{\sigma}$ is the perturbatively calculable partonic cross section, $f_{a,b}^{A,B}(x,Q^2)$ the parton distribution function (PDF) for a parton flavour a or b inside the colliding particles, covering the non-perturbative part of the cross section. The function $f(x,Q^2)$ describes the distribution of the momentum fraction x in the incident proton of a particular parton type at an energy scale Q^2.
The interpretation of measured cross sections in hadron collisions requires an adequate knowledge of the PDFs. For PDFs describing the partons inside a proton (or inside a neutron) there are several PDF sets based on measurements from e-p colliders, hadron colliders and fixed target experiments. Examples besides the previously mentioned MSTW08 set (*cf.* Chapter 1) are CT10 [183] or NNPDF [72].

Nuclear PDFs $f_a^A(x,Q^2)$ [3] are defined as linear combinations of the PDFs of the bound protons ($f_a^{p,A}(x,Q^2)$) and neutrons ($f_a^{n,A}(x,Q^2)$) as

$$f_a^A(x,Q^2) = \frac{N_Z}{N_A} f_a^{p,A}(x,Q^2) + \frac{N_A - N_Z}{N_A} f_a^{n,A}(x,Q^2). \tag{10.2}$$

The PDFs for the bound nucleons are usually described as

$$f_a^{i,A}(x,Q^2) = R_a^A(x,Q^2) f_a^i(x,Q^2), \tag{10.3}$$

where $f_a^i(x,Q^2)$ is the free nucleon baseline PDF and $R_a^A(x,Q^2)$ parametrises modifications due to nuclear effects and is called *nuclear modification factor*.

10.2.2 Nuclear Parton Distribution Function Sets

There are several nPDF sets or sets of nuclear modification factors available at next-to-leading order (NLO). The latest ones are EPS09 [141], HKN07 [170], DSSZ [127] and nCTEQ [181, 182, 215]. These sets differ among other points in the baseline free proton PDF that they use (CTEQ6 or MSTW08 and its predecessor MRST98). The data used in the PDF fits are from deep inelastic scattering (DIS) of leptons or protons on nuclei, in particular on deuteron. Also neutrino-nuclei DIS as well as deuteron-gold scattering at the Relativistic Heavy Ion Collider (RHIC) serve as input data in some sets. Table 10.1 summarises the status and properties of the different nPDF sets.

10.2.3 Properties of Nuclear Parton Distribution Functions

The nuclear modification factor $R_a^A(x,Q^2)$ shows – independently of the actual nPDF set – a distinct behaviour as a function of x (*cf.* Fig. 10.2): at small x-values, $R_a^A(x,Q^2)$ is lower than unity (*shadowing*). This effect can be interpreted as interference between scattering amplitudes with different numbers of nucleon interactions [156,165,178]. This effect appears at small x-values as the coherence length l_c in an interaction is typically

$$l_c = (2m_N x)^{-1}, \tag{10.4}$$

[3] x is in context of nuclei in the following also named x_A to distinguish it from the x-value of the parton in the bare proton or neutron.

Table 10.1: Summary of nPDF properties and data included in the PDF fits. (adapted from Refs [181] and [206])

Property	**nCTEQ**	**DSSZ**	**HKN07**	**EPS09**
Neutral current DIS $\ell + A/\ell + d$	✓	✓	✓	✓
Drell-Yan DIS $p + A/p + d$	✓	✓	✓	✓
RHIC π^0 $d + \mathrm{Au}/p + p$		✓		✓
ν-A DIS				✓
Q^2 cut in DIS	4 GeV^2	1.69 GeV^2	1 GeV^2	4 GeV^2
Baseline free proton PDF	CTEQ6M	MSTW08	MRST98	CTEQ6.1
Heavy quark treatment[a]	GM-VFNS	GM-VFNS	ZM-VFNS	ZM-VFNS

[a] ZM-VFNS: Zero Mass - Variable Flavour Number Scheme considers also c and b quarks as massless while the number of light quark flavours increases with scale if $\mu_F > m_{b,c}$.
GM-VFNS: General Mass - Variable Flavour Number Scheme considers the quarks as massive at low energy, but coincides with the ZM-VFNS at high scale for $\mu_F \gg m_b$. So it can be considered as a merging of ZM-VFNS and a Fixed Flavour Number Scheme (FFNS) where the later has the problem of treating only light quarks (u, d, s) and gluons as constituents of the nucleon.

where m_N is the mass of the nucleon. For small enough momentum fractions, the coherence length is much larger than the nucleon-nucleon distance inside a nucleus ($r_{NN} \approx 1.73$ fm) such that the interaction happens coherently with several nucleons [149].
$R_a^A(x, Q^2)$ is smaller than unity also for x-values between about 0.4 and 0.8, which corresponds to the *European Muon Collaboration (EMC) effect*. This effect has been first observed in DIS of muons on iron and deuteron nuclei by the European Muon Collaboration and is still not yet understood [68, 231].
At x-values close to one, $R_a^A(x, Q^2)$ is above unity as a result of the motion of nucleons inside their nucleus (*Fermi motion*).
Between the regimes of shadowing and the EMC-effect there is an other interval in x ($x \approx 0.1$) where $R_a^A(x, Q^2)$ is larger than unity. This effect is called *anti-shadowing*. It is not associated to any particular dynamical effect, but rather to the sum rule of the PDF [91].

Figure 10.2 also reveals two other features of nPDFs: firstly, they are poorly constrained, especially at small x-values due to a lack of high quality input data for their fits. Measurements of QCD processes in the forward direction, as they can be done at LHCb, would serve as very much appreciated input for nPDFs fits at small x-values, but also at x-values close to one, similar to the impact of the corresponding measurements with pp data in LHCb [70, 71].
Secondly, the gluon nPDF has particularly weak constraints at large values of x as most of the input data available so far come from Drell-Yan processes in pA fixed target experiments or DIS of electrons on nuclei.

A further aspect of nPDFs – not discussed in the following chapter – is that in contrast to the functions of bare nucleons they can show an impact-parameter dependance. This effect is a result of the scattering including a nucleus as an extended object and appears as a natural consequence from the modelling of the origin of shadowing (*e.g.* through the leading-twist approach as described in Ref. [149]). The effect is particularly pronounced in nucleus-nucleus scattering,

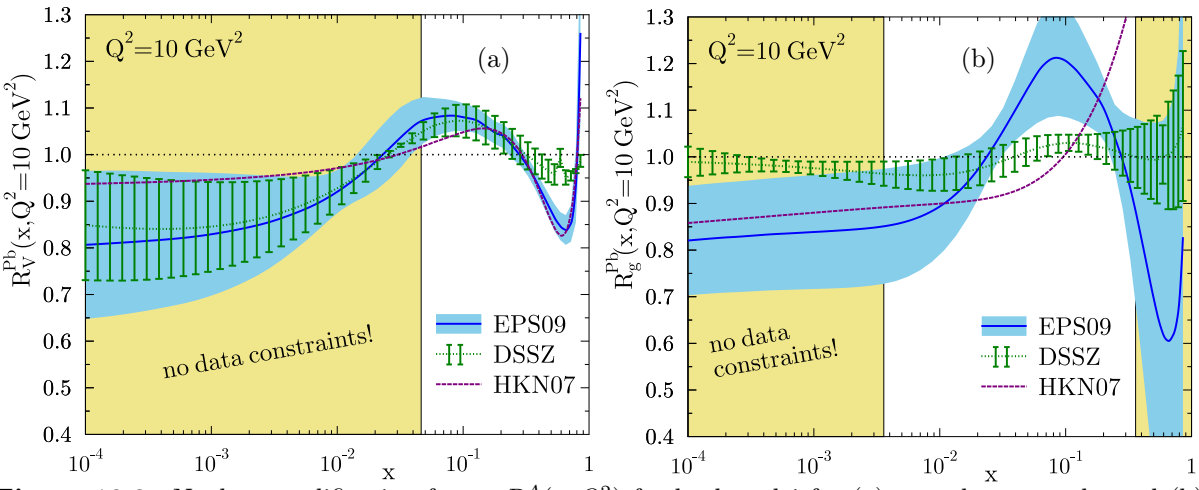


Figure 10.2: Nuclear modification factor $R_a^A(x, Q^2)$ for lead nuclei for (a) up valence quarks and (b) gluons from different nPDF sets (EPS09, DSSZ, HKN07) as a function of the fractional parton momentum x at $Q^2 = 10$ GeV^2. The yellow areas show the regions with no or very weak direct constraints from data. (from Ref. [206])

but should also modify proton-nucleus and lepton-nucleus scattering. So far, EPS09 is the only nPDF set taking this phenomenon into account [140].

10.3 Z Production in Hadron Collisions

10.3.1 Production Mechanisms

The production of Z bosons in hadron collisions happens at leading order through the Drell-Yan process in quark-anti-quark annihilation (*cf.* Fig. 10.3). Besides the production of on-shell Z bosons there is also an interference with off-shell photons, γ^*, produced through quark-anti-quark annihilation. These two processes are not distinguished in the following and considered combined for the region around the pole mass of the Z boson. Taking into account the factorisation theorem

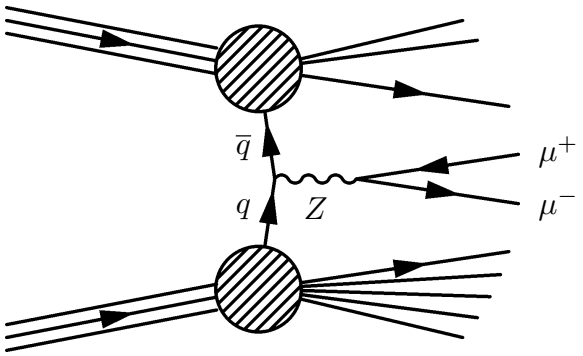


Figure 10.3: Feynman diagram for Z production through the Drell-Yan process which is the leading order contribution to Z production at hadron colliders.

(*cf.* Eq. (10.1)), the leading-order cross section for Z bosons decaying into opposite-charged muons can be expressed as

$$\sigma_{Z\to\mu^+\mu^-,A_1A_2} = N_{A_1}N_{A_2}\int \mathrm{d}x_1\mathrm{d}x_2 \sum_q f_q^{N_1}(x_1) f_{\overline{q}}^{N_2}(x_2)\sigma_{q\overline{q}\to Z\to\mu^+\mu^-}(x_1x_2\sqrt{s_{NN}}), \tag{10.5}$$

where $x_{1,2}$ are the momentum fractions of the annihilating quark q and anti-quark $\overline{q}$ inside the two colliding nuclei $N_{1,2}$ with $N_{A_{1,2}}$ nucleons and $\sqrt{s_{NN}}$ is the centre-of-mass energy per nucleon-nucleon pair.

Higher order processes can be produced involving gluons – also as initial state partons – and lead to additional jets of final state particles originating from the hadronisation of gluons or quarks emerging from the hard scatter. Figure 10.4 shows two examples of NLO Feynman diagrams for Z production.

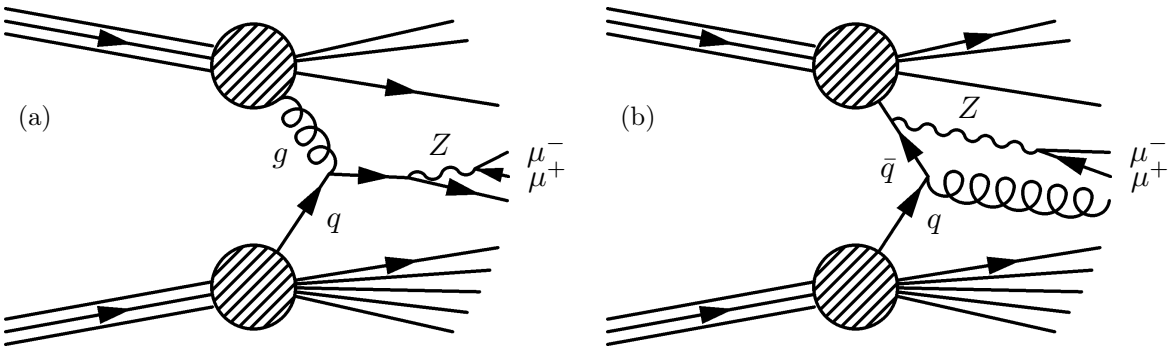


Figure 10.4: NLO Feynman diagrams for Z production in hadron colliders. Figure (a) shows the diagram for gluon-quark scattering resulting in a quark jet in the final state while Fig. (b) displays the diagram for quark-anti-quark scattering producing a gluon jet.

The relation between the energy scale, Q^2, of the hard scatter (for Z production $Q^2 = M_Z^2$), the centre of mass energy $\sqrt{s}$ and the fractional momenta, x_1 and x_2, of the two partons in the hard scatter is given by

$$Q^2 = x_1^2 x_2^2 s. \tag{10.6}$$

Further, the fractional momenta x_1 and x_2 and the rapidity y of the rest frame of the hard scatter in the lab frame are connected by

$$x_{1,2} = \frac{Q}{\sqrt{s}} e^{\pm y} \tag{10.7}$$

with the boundary conditions $x_{1,2} \leq 1$. So large absolute values in rapidity correspond to the situation where x_1 is high and x_2 is low or vice versa.
Due to its forward geometry, the LHCb experiment can therefore probe fractional momenta down to very small values as well as values close to one. This has been for example done with measurements of Z and W boson production in pp collisions [23, 30, 34, 189, 190] or Drell-Yan production of dimuon pairs at energies below the mass of the Z boson [188]. Figure 10.4 shows the region in x_A-Q^2 space that can be probed by LHCb with proton-lead collisions at $\sqrt{s_{NN}} = 5\,\mathrm{TeV}$.

For $Q^2 = M_Z^2$, x_A-values down to about 2×10^{-4} can be probed. This low x_A-range could be probed with central detectors such as CMS, ATLAS or ALICE only at energy scales comparable to the J/ψ mass with the disadvantage of requiring usage of the DGLAP evolution [53, 131, 162] to calculate the nPDFs for higher energy scales.

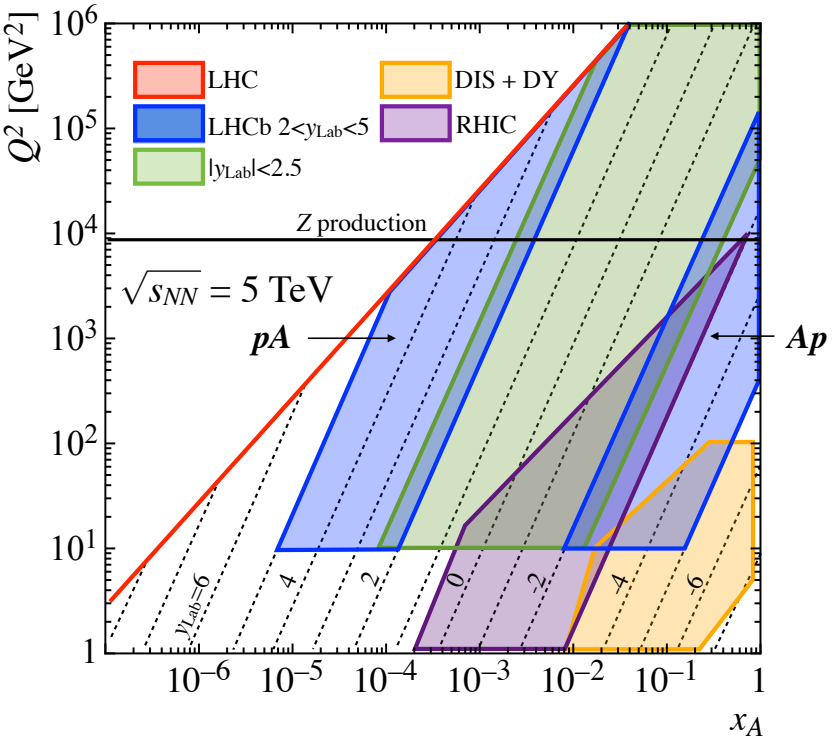


Figure 10.5: Regions in x_A-Q^2 that LHC, RHIC or fixed target experiments (DIS + DY) are sensitive to. The centre-of-mass energy per proton-nucleon pair for LHC is $\sqrt{s_{NN}} = 5\,\text{TeV}$. The red line shows the kinematic boundary for this particular centre-of-mass energy. The black horizontal line corresponds to the mass of the Z boson indicating that LHCb is sensitive down to $x_A \approx 2 \cdot 10^{-4}$ for measurements of the Z production. The green area shows the phase space corresponding approximately to the ATLAS and CMS detector and the central part of the ALICE detector. Process-dependent phase spaces of these experiments can be seen in Ref. [213]. The regions for RHIC and fixed target experiments have been also taken from Ref. [213]. The corresponding rapidity in the lab frame is shown by the dashed lines and indicated by the numbers in the plot.

10.3.2 Next-to-Next-to-Leading Order Predictions

The most precise predictions for the Z production cross section are based on NNLO calculations, using for example FEWZ [151, 191]. Table 10.2 shows these predictions for pp collisions in the fiducial volume of LHCb measurements ($2 < \eta_\mu < 4.5$, $p_{\text{T},\mu} > 20\,\text{GeV}/c$, $60 < m_{\mu^+\mu^-} < 120\,\text{GeV}/c^2$) for pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 5\,\text{TeV}$, the latter with beam energies of 4 TeV and 1.58 TeV corresponding to the conditions in the recorded proton-nucleus collisions.
The nuclear effects can be incorporated by describing their impact using the EPS09 nPDF set at NLO. The corresponding values are also shown in Tab. 10.2. From a simple scaling of the cross

Table 10.2: Predictions for the Z cross section measured in LHCb at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 5\,\text{TeV}$ using FEWZ at NNLO and different PDF sets. In the case of 5 TeV also predictions with nuclear modifications based on the EPS09 nPDF set evaluated at NLO are shown. The first uncertainty is due to the scale uncertainty – evaluated by changing the renormalisation and factorisation scales by a factor two up and down –, the second one due to the uncertainty on the PDF set and the third one, if displayed, due to the uncertainty on the nPDF set.

$\sigma_{Z\to\mu^+\mu^-}$	**Value [pb]**
DYNNLO $\sqrt{s} = 7$ TeV [23]	
MSTW08	$74.3^{+1.0}_{-1.5}(\text{theo.})^{+1.6}_{-1.4}(\text{PDF})$
ABKM09	$77.2^{+1.0}_{-1.5}(\text{theo.})^{+0.7}_{-0.7}(\text{PDF})$
JR09	$73.7^{+1.0}_{-1.5}(\text{theo.})^{+1.2}_{-0.9}(\text{PDF})$
NNPDF	$76.5^{+1.0}_{-1.5}(\text{theo.})^{+1.1}_{-1.1}(\text{PDF})$
HERA15	$74.9^{+1.0}_{-1.5}(\text{theo.})^{+1.2}_{-1.8}(\text{PDF})$
CTEQ6M (NLO)	$72.3^{+1.7}_{-2.3}(\text{theo.})^{+1.7}_{-1.5}(\text{PDF})$
FEWZ $\sqrt{s} = 5$ TeV (4+1.58 TeV) no nucl. modifications	
MSTW08 pA configuration	$69.6^{+0.6}_{-0.4}(\text{theo.})^{+1.4}_{-1.3}(\text{PDF})$
MSTW08 Ap configuration	$13.5^{+0.1}_{-0.1}(\text{theo.})^{+0.3}_{-0.3}(\text{PDF})$
FEWZ $\sqrt{s} = 5$ TeV (4+1.58 TeV) + EPS NLO	
MSTW08 pA configuration	$63.1^{+0.5}_{-0.4}(\text{theo.})^{+1.3}_{-1.2}(\text{PDF})^{+0.1}_{-0.2}(\text{nPDF})$
MSTW08 Ap configuration	$12.6^{+0.1}_{-0.1}(\text{theo.})^{+0.3}_{-0.3}(\text{PDF})^{+0.2}_{-0.4}(\text{nPDF})$

section at $\sqrt{s} = 5\,\text{TeV}$ by the number of nucleons $N_A = 208$ in lead nuclei

$$\sigma_{Z\to\mu^+\mu^-,pA/Ap} = N_A \cdot \sigma_{Z\to\mu^+\mu^-,pp}, \tag{10.8}$$

an estimate for the cross section in proton-nucleus collisions, $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$, can be obtained. For the two beam configurations in LHCb the corresponding numbers are

$$\begin{aligned}\sigma_{Z\to\mu^+\mu^-,pA} &= 14.48^{+0.12}_{-0.09}(\text{theo.})^{+0.30}_{-0.27}(\text{PDF})\ \text{nb}\\ \sigma_{Z\to\mu^+\mu^-,Ap} &= \ \ 2.81^{+0.03}_{-0.03}(\text{theo.})^{+0.07}_{-0.06}(\text{PDF})\ \text{nb}\end{aligned}$$

and

$$\begin{aligned}\sigma_{Z\to\mu^+\mu^-,pA} &= 13.12^{+0.11}_{-0.08}(\text{theo.})^{+0.27}_{-0.24}(\text{PDF})^{+0.03}_{-0.10}(\text{nPDF})\ \text{nb}\\ \sigma_{Z\to\mu^+\mu^-,Ap} &= \ \ 2.61^{+0.03}_{-0.03}(\text{theo.})^{+0.07}_{-0.06}(\text{PDF})^{+0.03}_{-0.08}(\text{nPDF})\ \text{nb}\end{aligned}$$

including the nuclear modifications based on the EPS09 nPDF set at NLO.
Both cross sections are suppressed when including nuclear effects. But while the cross section for the pA configuration is suppressed by the nuclear shadowing, it is suppressed by the EMC-effect in the Ap configuration.
A further possible way to incorporate these nuclear effects, based on the leading-twist approach for nuclear shadowing phenomena [149], has been shown in Ref. [165]. The calculations therein for the differential Z production cross section $\mathrm{d}\sigma/\mathrm{d}y$ have been performed using the Collins-Soper-Sterman resummation formalism [119–121] implemented in the RESBOS computer code [69, 185]. These predictions show a suppression of the cross section due to the shadowing for Z rapidities $y > 0$ in the pA configuration. The suppression can be as large as 15 % for $y \approx 3.5$. For the Ap configuration, the effect of anti-shadowing leads to an enhancement of Z production at rapidities between 0.5 and 3.0. The maximal enhancement happens at $y \approx -2.0$ and is about 5 %. Figure 10.6 shows the predicted ratio between the cross section with and without nuclear effects for centre-of-mass energies per nucleon pairs of 4.4 and 8.8 TeV, corresponding to a proton beam energy of 3.5 and 7 TeV, respectively.
A particularly sensitive quantity to observe nuclear effects is – due to this behaviour of enhancement and suppression of the cross section – the forward-backward ratio of the cross sections

$$R_{\text{FB}}(\sqrt{s_{NN}}, |y|) = \frac{\sigma_{Z\to\mu^+\mu^-}(\sqrt{s_{NN}}, +|y|)}{\sigma_{Z\to\mu^+\mu^-}(\sqrt{s_{NN}}, -|y|)}. \tag{10.9}$$

Following the above argumentation, R_{FB} is expected to be smaller than one in the overlapping interval of the rapidity y in the centre-of-mass frame $2.5 < |y| < 4.0$ of both beam configurations.

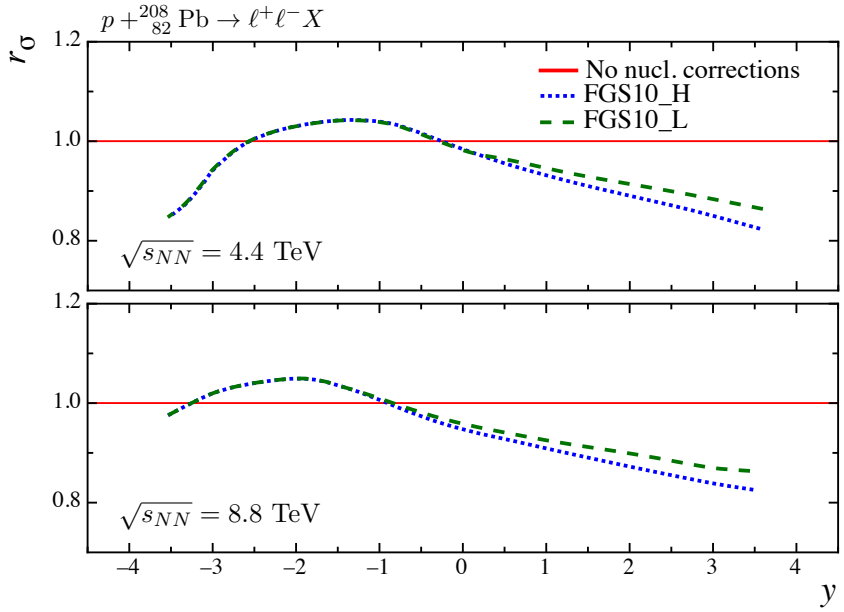


Figure 10.6: The nuclear correction factor r_σ on the Z production cross section as a function of y calculated from the leading-twist approach for nuclear shadowing (FGS10 [149]) and the Collins-Soper-Sterman resummation formalism. The top plot shows the situation for a centre-of-mass energy of $\sqrt{s_{NN}} = 4.4\,\mathrm{TeV}$, the bottom one for $\sqrt{s_{NN}} = 8.8\,\mathrm{TeV}$. FGS10_H and FGS10_L show the upper and lower limit on the nuclear shadowing. (adapted from Ref. [165])

11

Z Production in Proton-Lead Collisions at LHCb

This chapter describes the measurement of the $Z \to \mu^+\mu^-$ production cross section $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$ in the forward and backward direction at LHCb in proton-lead collisions. The measurement is based on the data sample corresponding to a total integrated luminosity of 1.6 nb^{-1} collected at the beginning of 2013. The cross section is estimated as

$$\sigma_{Z\to\mu^+\mu^-,pA/Ap} = \frac{N_{\text{cand}} \cdot \rho}{\mathcal{L} \cdot \varepsilon^{\text{GEC}} \cdot \varepsilon^{\text{REC}} \cdot \varepsilon^{\text{SEL|REC}} \cdot \varepsilon^{\text{TRIG|SEL}} \cdot \varepsilon^{\text{MUON|TRIG}}}, \tag{11.1}$$

where N_{cand} is the number of reconstruced candidates, ρ the purity (*i.e.* the fraction of signal events) and $\mathcal{L}$ the integrated luminosity; ε^{GEC} is the efficiency of the Global Event Cuts (GEC)[1] described in the first section. ε^{REC} is the reconstruction efficiency, $\varepsilon^{\text{SEL|REC}}$ the selection efficiency for reconstructed candidates, $\varepsilon^{\text{TRIG|SEL}}$ the selection efficiency for selected candidates and $\varepsilon^{\text{MUON|TRIG}}$ the muon-identification efficiency for triggered candidates, respectively.

The data samples used in this analysis are described in Sec. 11.1 while Sec. 11.2 describes the procedure to deal with the different track multiplicities in pA, Ap and pp data. Section 11.3 describes the estimation of the purity ρ, Sec. 11.4 describes the determination of the different efficiencies, Sec. 11.5 the estimation of the associated systematic uncertainties and Sec. 11.6 describes the calculation of the final result.

11.1 Data Samples

11.1.1 General Properties

The data samples used in this analysis were collected during the 2013 pA and Ap runs. For both beam configurations, data have been taken with both LHCb magnet polarities. Table 11.1 shows

[1] The GEC have the aim to reject at the trigger stage events that have a very high multiplicity and therefore consume a large amount of computing resources.

the details of the samples while Fig. 10.1 shows the integrated luminosity as a function of time. In the pA and Ap runs, the L0 and HLT2 stage of the trigger were set in pass-through mode. Besides the change in the trigger configuration with respect to the pp collisions also the GEC have been changed with respect to the cuts applied in pp collisions to take the different multiplicities in pPb collisions into account. The GEC for 2012 data (pp collisions) have been

- Number of SPD hits < 600 (Muon), 900 (Dimuon) (L0);
- Number of IT clusters < 3000 (HLT1);
- Number of OT clusters $< 15'000$ (HLT1);
- Number of VELO clusters $< 10'000$ (HLT1);
- Number of VELO tracks < 350 (HLT2).

The relevant cuts in proton-lead collisions are

- Number of OT clusters $< 15'000$ (HLT1);
- Number of IT clusters < 3000 (HLT1);
- Number of VELO clusters < 6000 (HLT1);

and are the same for both beam configurations.
The data collected in four runs[2] (Run numbers: 135710, 135711, 135900, 135901) have been discarded due to inefficiencies in the Muon System. Also data taken with specific trigger configurations (Trigger configuration keys: 0x6a1710, 0x6e0049, 0x6f004a) are not taken into account[3]. The analysed data correspond in total to $\mathcal{L} = 1.099 \pm 0.021$ nb^{-1} for pA and 0.521 ± 0.011 nb^{-1} for Ap. The integrated luminosity has been calibrated by Van der Meer scans for each beam configuration separately [99, 226].
The data sample of 2 fb^{-1} collected in pp collisions at $\sqrt{s} = 8$ TeV is used to estimate various quantities (*e.g.* purity, trigger efficiencies) that cannot be estimated directly from the pA and Ap data due to their low statistics.

11.1.2 Candidate Selection

The Z candidates are selected using two opposite-charged tracks identified as muons with a high transverse momentum p_{T}. The detailed selection cuts are listed in Tab. 11.2. The selection criteria on the muons in terms of η and p_{T}, which are the same as in the $Z \to \mu^+\mu^-$ production measurement in pp collisions [190], are motivated by the acceptance of LHCb ($\eta > 2.0$), the kinematic cut-off ($\eta < 4.5$) and the mass of the Z boson ($p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$; $2 \cdot p_{\mathrm{T}} < m_Z$). Figure 11.1 shows the p_{T} and η distributions of the muons in Z candidates from simulation.

Furthermore, it is required that at least one of the two muons from the Z candidates in the pA and

[2] An LHC fill, which corresponds to the whole data taken with the same injected beams, is separated into several runs to facilitate the offline data quality management.
[3] Data from 0x6a1710 are removed due to inefficiencies in the Muon System while the data with the other two keys are not taken into account due to slightly different L0 and prescale configurations.

Table 11.1: Summary of the data and MC samples used in the Z production measurement in proton-lead collisions.

Sample	$\sqrt{s}$ **[TeV]**	**(MC)/Reco/Stripping version**	**Sample size**[a]
$Z \to \mu^+\mu^-$ pA	5	R14/S20r3	1.099 nb^{-1}
$Z \to \mu^+\mu^-$ Ap	5	R14/S20r3	0.521 nb^{-1}
$Z \to \mu^+\mu^-$ pp	8	R14/S20	1.97 fb^{-1}
$Z \to \mu^+\mu^-$ pp MC	8	2012 (Sim06b/R14/S20)	1'000'000
$J/\psi \to \mu^+\mu^-$ pA	5	R14/S20r3	1.099 nb^{-1}
$J/\psi \to \mu^+\mu^-$ Ap	5	R14/S20r3	0.521 nb^{-1}
$J/\psi \to \mu^+\mu^-$ pp	8	R14/S20	3.33 pb^{-1}
$J/\psi \to \mu^+\mu^-$ pp MC	8	2012 (Sim08a/R14a/S20)	20'000'000
$b\bar{b} \to \mu^+\mu^- X$ pp MC	7	MC11a (Sim05a/R12/S17)	5'000'000

[a] For data, the number shows the total integrated luminosity; for MC, the number of generated events in the sample.

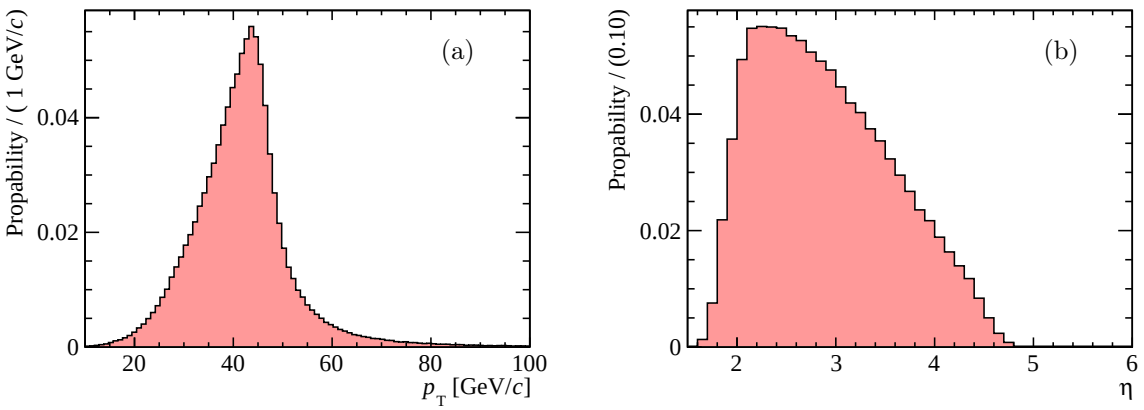


Figure 11.1: The (a) p_{T} and (b) η distribution of muons in $Z \to \mu^+\mu^-$ MC in the LHCb acceptance.

Ap sample is triggered by the Hlt1SingleMuonHighPT line requiring $p > 8\,\text{GeV}/c$, $p_\text{T} > 4.8\,\text{GeV}/c$ and $\chi^2_\text{trk}/\text{ndf} < 4$.
As mentioned above, beside a sample of $Z \to \mu^+\mu^-$ candidates also a sample of $J/\psi \to \mu^+\mu^-$ candidates and a sample with the same selection criteria as for the $Z \to \mu^+\mu^-$ candidates but the same charge of the two muon candidates are selected[4].

Table 11.2: Description of the selection cuts for $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ candidates in pA/Ap.

Variable	$\boldsymbol{Z \to \mu^+\mu^-}$	$\boldsymbol{J/\psi \to \mu^+\mu^-}$
Muons		
p_T [GeV/c]	> 20[a]	> 2.5
η	$[2.0, 4.5]$	$[2.0, 4.5]$
IP_PV [mm]	< 0.1	< 0.1
E/pc [b]	< 0.5	< 0.5
$\text{Prob}(\chi^2_\text{trk}, \text{ndf})$[c]	> 0.01	> 0.01
IsMuon	yes[d]	yes
$\Delta\text{LL}_{\mu-\pi}$	–	> 0
$\boldsymbol{Z/J/\psi}$ **candidate**		
$\chi^2_\text{vtx}/\text{ndf}$	–	< 25
$m_{\mu^+\mu^-}$ [GeV/c^2]	$[60, 120]$[a]	$[3.01, 3.17]$

[a] For the evaluation of the purity, a larger window for $m_{\mu^+\mu^-}$ of $[40, 160]\,\text{GeV}/c^2$ and a looser cut of $p_\text{T} > 15\,\text{GeV}/c$ (due to the lower minimal mass) are applied.

[b] E is the sum of the energy measured in the electromagnetic and hadronic calorimeter associated to the muon track.

[c] $\text{Prob}(\chi^2_\text{trk}, \text{ndf})$ is the χ^2-probability from the muon track fit.

[d] For the evaluation of the muon-identification efficiency a dedicated sample where only one of the two muon needs to fulfil the IsMuon requirement is selected. The IsMuon requirement is described in Sec. 3.4.1

As a higher background contamination due to the higher multiplicity, especially in the *Ap* collisions, is expected, additional cuts with respect to the $Z \to \mu^+\mu^-$ production measurement in *pp* collisions (*cf.* Ref. [189]) are applied. These cuts use the impact parameters (IP_PV) of the two muons with respect to the associated primary vertex, the ratio between the sum of the energy in the electromagnetic and hadronic calorimeter associated to the muons and the momentum measured by the tracking system (E/pc) as well as the χ^2-probability from the muon track fit. Their goal is to reject backgrounds from semileptonic heavy quark decays and tracks from charged hadrons mis-identified as muons from decays in flight – pions or kaons decay leptonically into a muon and a neutrino – or from punch-throughs into the Muon System[5].
The distributions of the three variables for *pp* and *pA* or *Ap* data, $Z \to \mu^+\mu^-$ MC as well as background from simulation and from minimal bias data are shown in Fig. 11.4. Figure 11.2 shows

[4] The selection has been done for all samples with DaVinci v33r5 using database tags `cond-20131028` and `dddb-20130503-1` for data and `Sim08-20130503-1-vc-mu/md100` and `Sim08-20130503-1` for 2012 MC as well as `sim-20130503-vc-mu/md100` and `MC11-20111102` for MC11a.

[5] Punch-through refers to hadrons having sufficient energy to transverse the calorimeter system and are detected by the muon stations M2-5.

that most of the background considering a looser invariant dimuon mass cut ($m_{\mu^+\mu^-} > 40\,\text{GeV}/c^2$) have larger muon IPs than Z signal candidates.

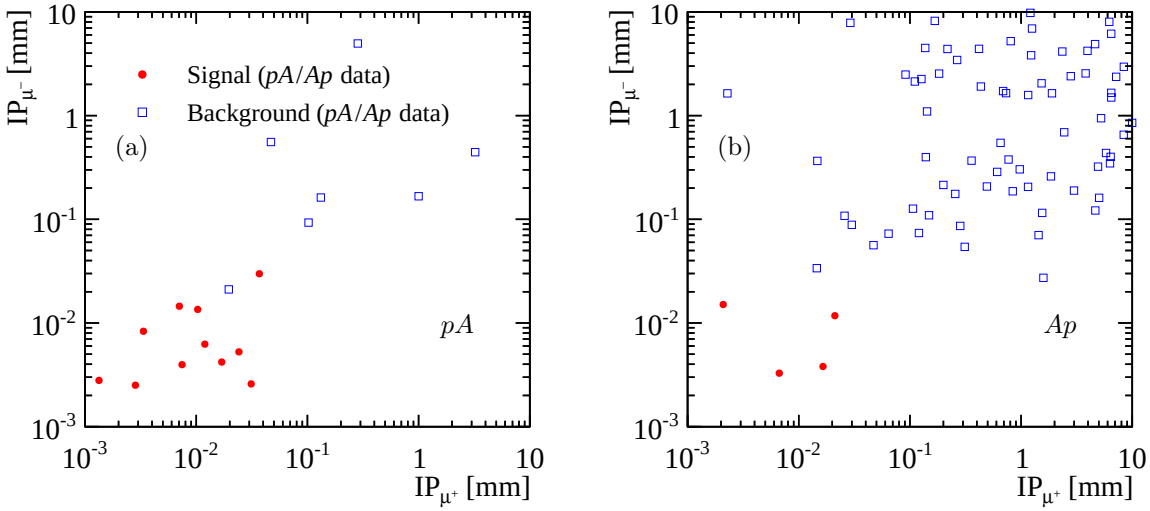


Figure 11.2: Two-dimensional distributions of the muon IP with respect to the PV for (a) pA and (b) Ap data. The signal candidates have been selected in the usual way while the background candidates fulfil the same criteria as the signal candidates except lying outside the invariant dimuon mass window and/or not fulfilling the additional cuts on E/pc, IP_{PV} and $\text{Prob}(\chi^2_{\text{trk}},\text{ndf})$.

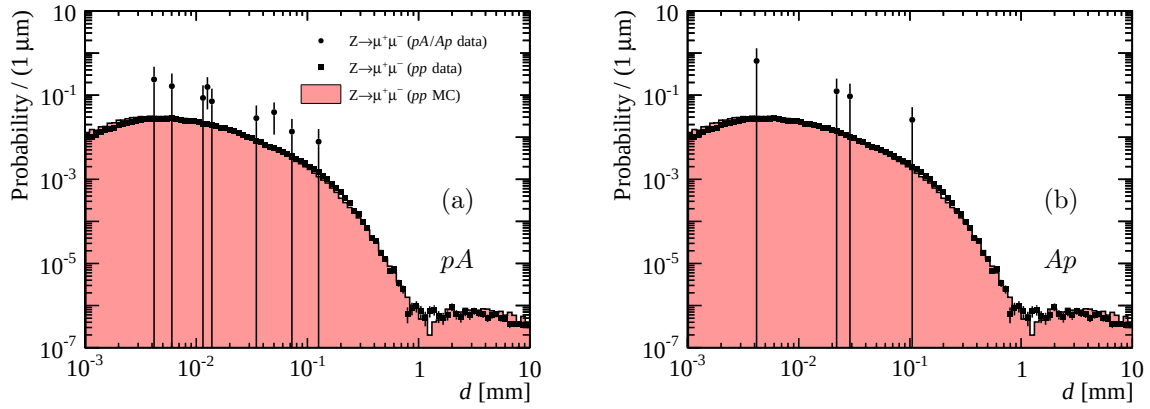


Figure 11.3: Distribution of the distance between the dimuon vertex and the closest primary vertex for (a) pA and (b) Ap data compared to pp data and MC.

Figure 11.5 shows the invariant dimuon mass and rapidity distribution of the $Z \to \mu^+\mu^-$ candidates in pA and Ap data, respectively. The observed yields are 11 candidates for pA and 4 candidates for Ap.

As a check of the background contamination of the selected Z candidates, their distribution of the distance between the dimuon vertex and the closest primary vertex has been studied (*cf.* Fig. 11.3). From comparing these distributions with the same for Z candidates in pp and MC which have a purity of almost 100 %, there is no visible background contamination present in the

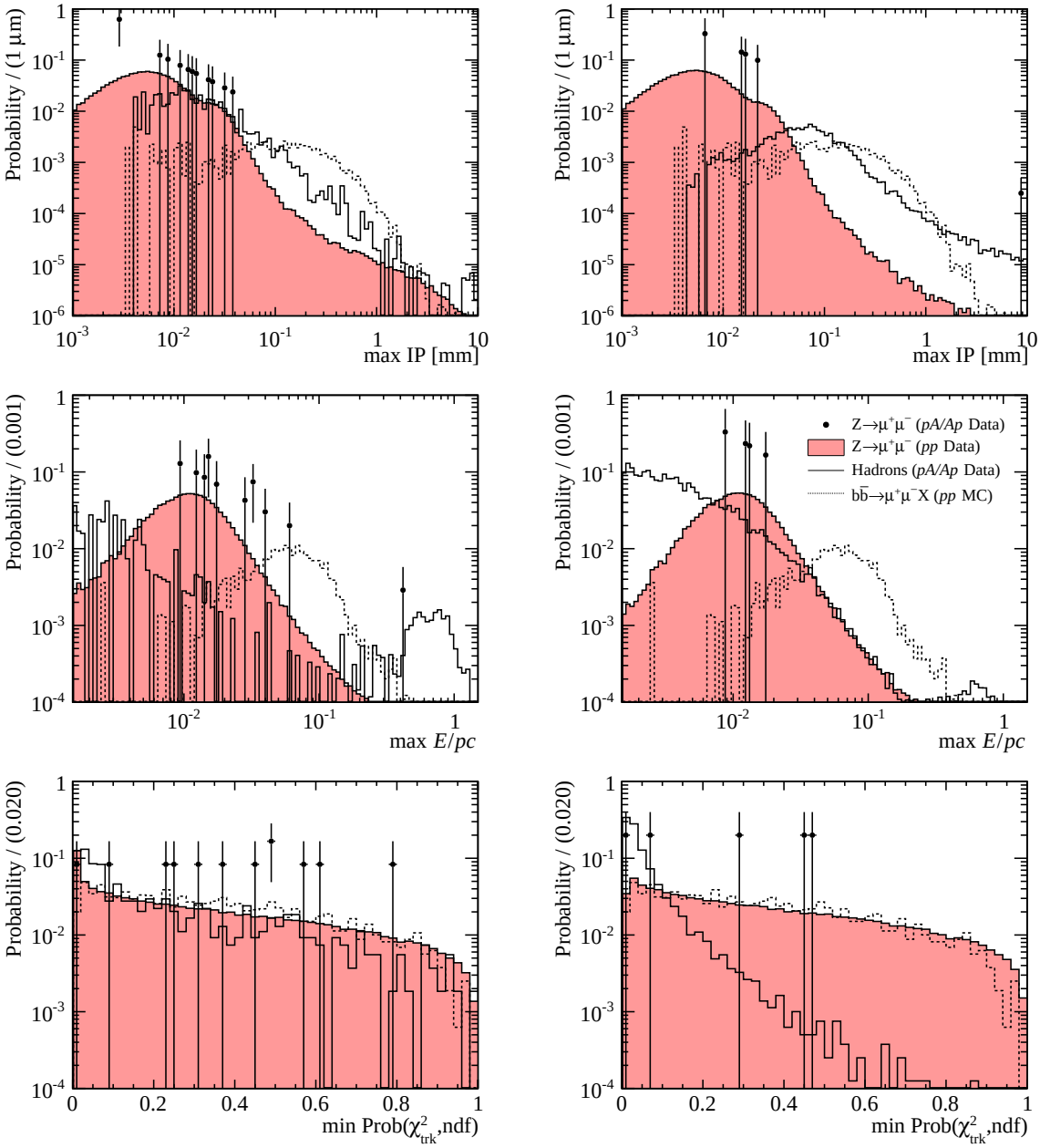


Figure 11.4: The distributions of (top row) the larger $\mathrm{IP}_{\mathrm{PV}}$, (center row) the larger E/pc value and (bottom row) the smaller $\mathrm{Prob}(\chi^2_{\mathrm{trk}}, \mathrm{ndf})$ of the two muon candidates. The plots show the distributions for $Z \to \mu^+\mu^-$ candidates in (left column) pA and (right column) Ap by the black dots with error bars. The black solid lines show the distributions for $Z \to \mu^+\mu^-$ candidates without any muon-id requirement in pA and Ap. For comparison, also the distributions for $b\bar{b} \to \mu^+\mu^- X$ from simulated pp collisions (black dashed line) and $Z \to \mu^+\mu^-$ candidates from pp data (red filled histogram) are shown.

Z candidates from proton-lead data.

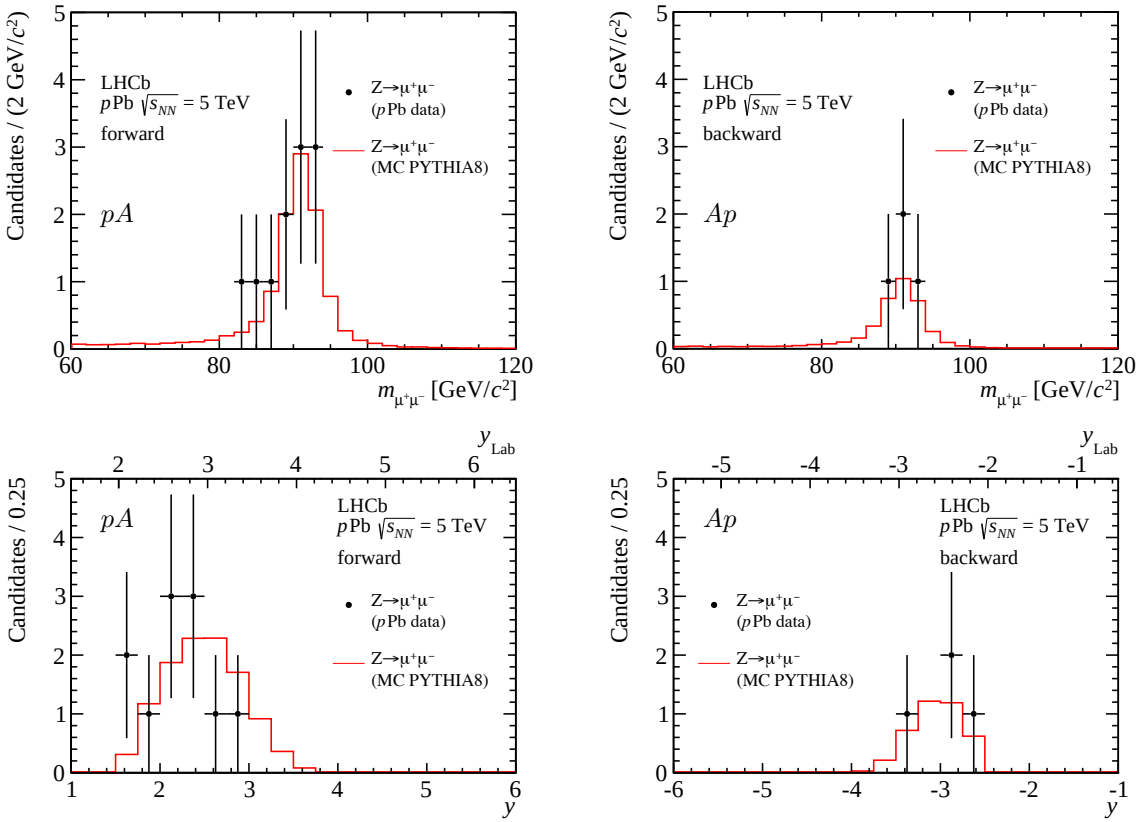


Figure 11.5: The invariant dimuon mass distribution (top row) and rapidity distribution (center row) of the $Z \to \mu^+\mu^-$ candidates in (left) pA (forward) and (right) Ap (backward) data. The horizontal axis in the rapidity plots show the Z rapidity in the laboratory frame, y_{Lab}, and in the centre-of-mass frame, y. As comparison the red lines in the plots show also the distributions from PYTHIA8 simulation using pp and pn collisions mimicking proton-lead collisions normalised to the number of observed candidates.

11.2 Multiplicity Reweighting

A major difference between the pp data and data from proton-lead collisions is the track multiplicity. As the pile-up in proton-lead collisions have been one and thus smaller than in proton-proton collisions, there is an only slightly higher track multiplicity in pA than in pp data. On the other hand, the multiplicity in Ap data is higher than in pp due to the colliding nucleus flying from the interaction region towards the Muon System. These differences are taken into account when using pp data to estimate quantities (*e.g.* purity, efficiencies) in the determination of $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$. The difference in track multiplicity is estimated from the $J/\psi \to \mu^+\mu^-$ samples. The distribution for the number of tracks in signal events is generated by applying the ${}_s\mathcal{P}lot$ technique [208] using the invariant dimuon mass $m_{\mu^+\mu^-}$ to discriminate signal and background (*cf.* Fig. 11.6). As signal model the sum of two Crystal Ball functions (${}_d\mathcal{CB}$, *cf.* Eq. (7.2)), where the tail parameters are fixed from MC (*cf.* Tab. 8.5), is used while the background distribution for $m_{\mu^+\mu^-}$

is described by an exponential distribution. The number of tracks is defined as the number of tracks reconstructed by any type of reconstruction algorithm in LHCb, but after removing all duplicates (*e.g.* a track from the forward tracking algorithm that has been also reconstructed by the track matching or as a downward track corresponding to a long track). Figure 11.8 shows that there is a strong correlation between the number of tracks and other variables representing the multiplicity such as the number of long tracks or the number of hits in the SPD.
The multiplicity reweighting is performed by using the ratio of the track multiplicity distributions in J/ψ events between pA or Ap and pp, as shown in Fig. 11.7, as weighting factor.

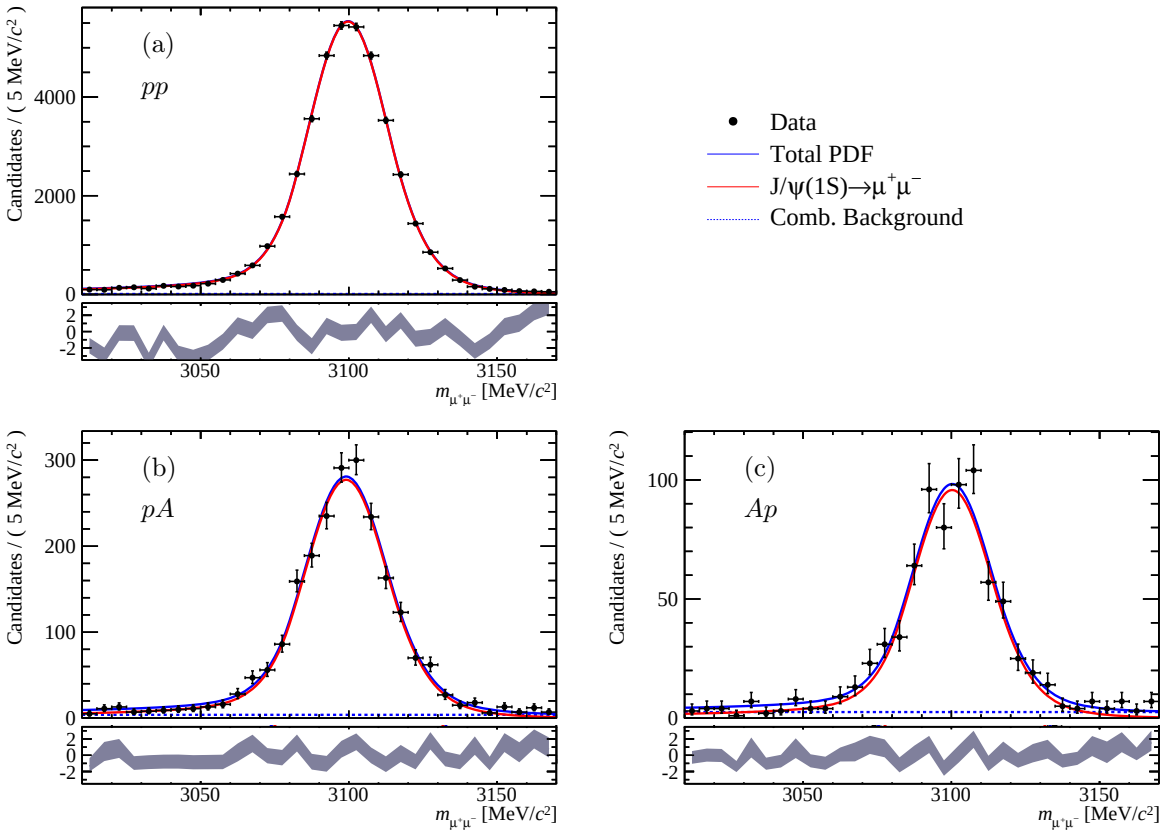


Figure 11.6: The invariant dimuon mass distribution of $J/\psi \to \mu^+\mu^-$ candidates in (a) pp, (b) pA and (c) Ap data. The distributions are fit with the sum of two Crystal Ball functions for signal (red solid line) and an exponential function for background (blue dashed line). The full fit model is shown by the solid blue line. The results of these fits are used to obtain the background subtracted track multiplicity distribution by the ${}_s\mathcal{P}lot$ technique.

A check for a systematic bias due to the choice of the variable describing the multiplicity has been performed. As alternative, the number of SPD hits has been used to describe the multiplicity with very similar results. Figure 11.10 shows the impact of the reweighting on the number of long tracks, SPD hits and TT clusters as different quantities reflecting the multiplicity. The reweighting is not perfect first of all due to GEC in the pp sample on the number of SPD hits and TT clusters, but also there is a region in the track multiplicity where there is no statistics in either the proton-lead or the proton-proton data and a reweighting factor of one has been

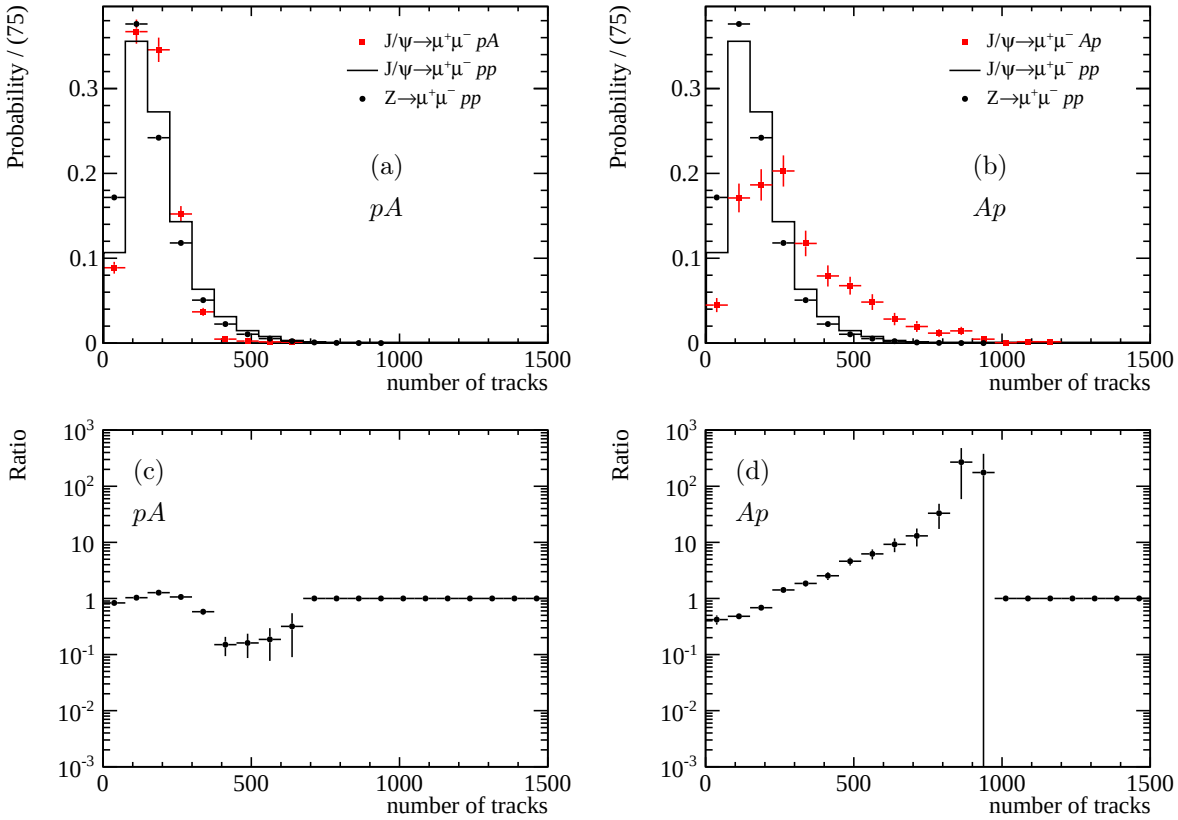


Figure 11.7: The top row shows the track multiplicity distributions in $J/\psi \to \mu^+\mu^-$ in (a) pA and (b) Ap data together with the distributions in $J/\psi \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-$ in pp data. The bottom row shows the ratio of the track multiplicity distributions for J/ψ between (c) pA and pp as well as (d) Ap and pp.

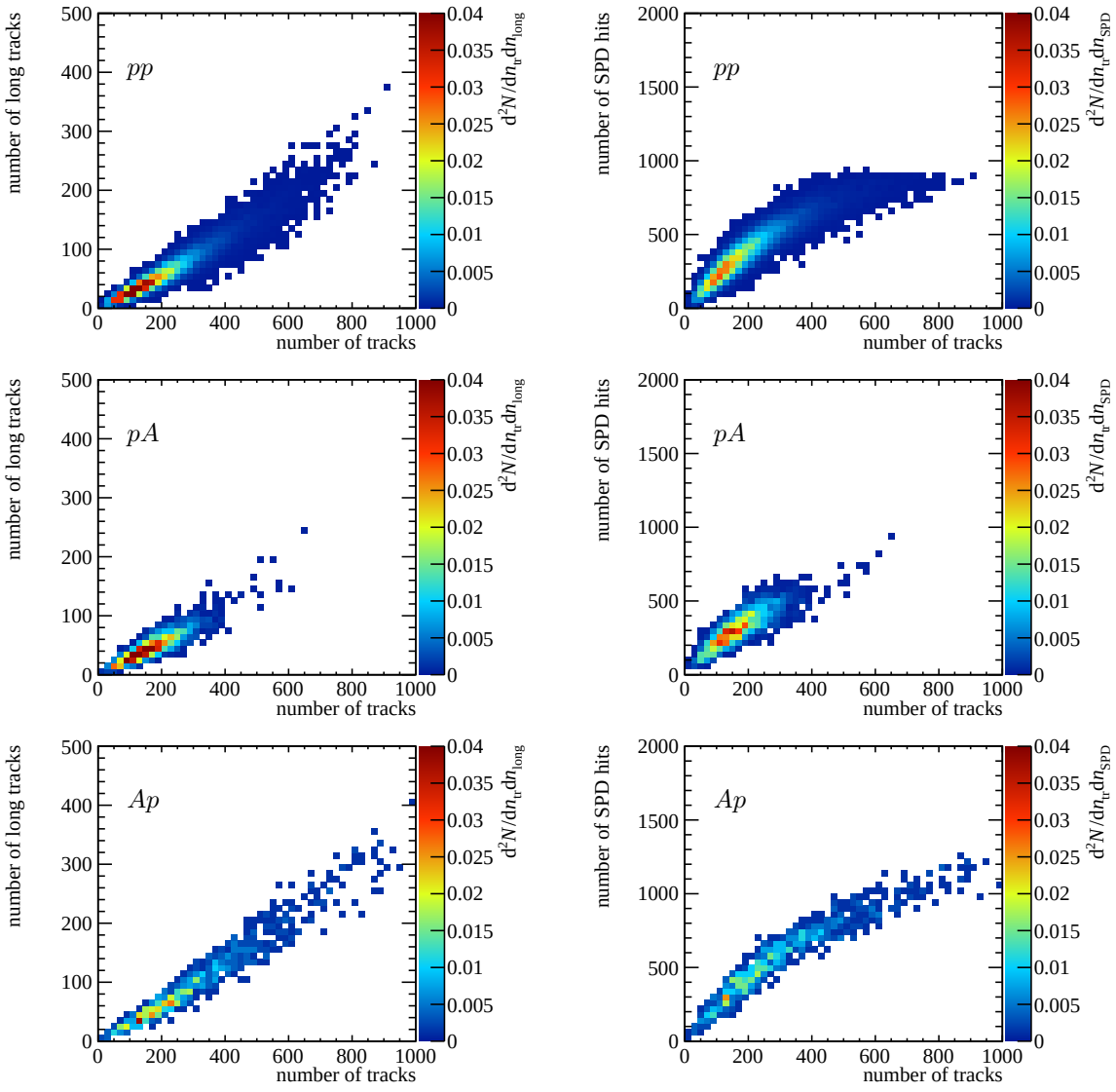


Figure 11.8: The correlation between the number of tracks and the number of (left column) long tracks and (right column) hits in the SPD for J/ψ candidates in (top row) pp, (center row) pA and (bottom row) Ap collisions. There is a strong correlation among the variables. There are a GEC that the number of SPD hits must be smaller than 600 in the single muon trigger and one of 900 in the dimuon trigger in proton-proton data (*cf.* Fig. (b)), which are not present in proton-lead data.

assumed.
Figure 11.9 shows the track multiplicity distribution of the events containing Z candidates in pA and Ap data. There is no significant deviation from the distributions of events containing J/ψ candidates shown in Fig. 11.7.

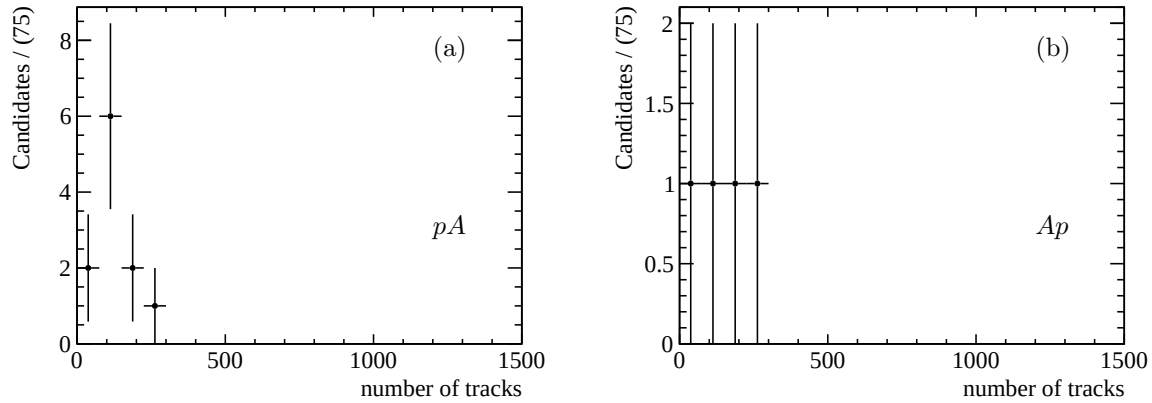


Figure 11.9: The track multiplicity distributions of events containing Z candidates in (a) pA and (b) Ap data.

The systematic uncertainty on the track multiplicity reweighting is estimated by recalculating the quantities that rely on the reweighting (purity as well as tracking, selection, trigger, muon-identification efficiency) without it and take the relative difference of $\rho/(\varepsilon^{\mathrm{REC}} \cdot \varepsilon^{\mathrm{SEL|REC}} \cdot \varepsilon^{\mathrm{TRIG|SEL}} \cdot \varepsilon^{\mathrm{MUON|TRIG}})$ between the calculation with and without reweighting as systematic uncertainty. The relative uncertainty is 1.5 % for pA and 2.0 % for Ap.

11.3 Purity Estimation

There are two relevant sources for background in the analysed samples of $Z \to \mu^+\mu^-$ candidates. The first one are candidates where one of the track is a hadron mis-identified as a muon, either by a decay in flight of a kaon or a pion or a punch-through of the particle into the Muon System (mis-identified hadron). The second source of background is from $b\bar{b}$ or $c\bar{c}$ pairs where both quarks decay semileptonicly into two muons (heavy quark). The contributions from other sources, such as $t\bar{t}$, $Z \to \tau^+\tau^-$ or W^+W^-, have been found so small in the measurement of Z production in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ [189] that they have been neglected in the analysis at hand.

11.3.1 Mis-identified Hadron

The background from mis-identified hadrons, which form together with an additional muon or with an other mis-identified hadron a $Z \to \mu^+\mu^-$ candidate should contribute equally to $Z \to \mu^+\mu^-$ candidates and to combinations of two muons with the same sign (same-sign candidates). Therefore it is estimated considering a sample of same-sign candidates in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ that pass the usual selection cuts except a lower p_{T} cut on the two muon candidates of $15\,\mathrm{GeV}/c$ and comparing it to a sample of opposite-sign candidates in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$.

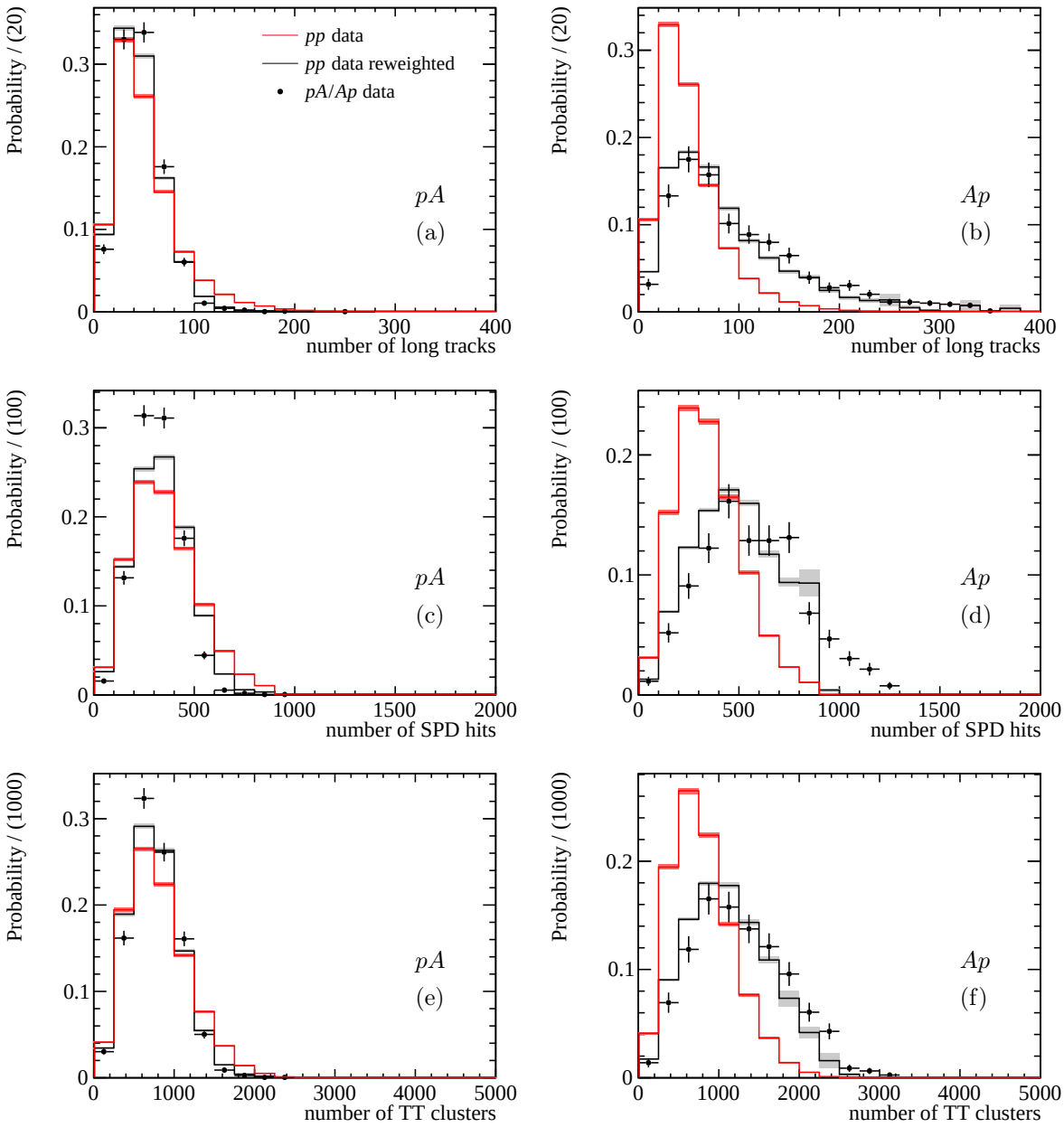


Figure 11.10: Impact of the track multiplicity reweighting on variables depending on the event multiplicity for (left column) pA and (right column) Ap. The top row shows the distribution of the number of long tracks, the center one of the number of SPD hits and the bottom row of the number of TT clusters.

The invariant dimuon mass range considered is between 40 and 160 GeV/c^2. Figures 11.11(c) and (d) show the invariant mass distribution of the same-sign candidates reweighted for the track multiplicity and normalised by the total number of $Z \to \mu^+\mu^-$ candidates (*cf.* Figs 11.11(a) and (b)) in the signal mass range between 60 and 120 GeV/c^2. The distribution is fit by an exponential function. The integral of this function in the signal mass range gives the fraction of background from mis-identified hadrons to the total number of candidates, which is for pA $(1.45 \pm 0.04) \times 10^{-3}$ and for Ap $(1.87 \pm 0.04) \times 10^{-3}$.
Figures 11.11(c) and (d) show at high dimuon masses a behaviour in data that is not well described by an exponential distribution. This feature is also present in the unweighted pp data (*cf.* Fig. 11.12(a)). As an alternative the sum of two exponential distributions has been fit to the data and used to extract the expected fraction of background candidates from the integral between 60 and 120 GeV/c^2 (*cf.* Figs 11.12(b) and (c)). The resulting fraction is in both cases (pA and Ap) smaller by only 0.20×10^{-3}. Thus the single exponential distribution is suitable enough for this purpose.

11.3.2 Heavy Quark Background

The amount of background from heavy quark decays is estimated with two different methods. Both methods use the sample of $Z \to \mu^+\mu^-$ candidates in pp collisions at $\sqrt{s} = 8$ TeV, selected by the usual criteria except a lower cut on the p_{T} of the muons of 15 GeV/c and a larger invariant dimuon mass range from 40 to 160 GeV/c^2.
The first method uses a cut of $\chi^2_{\mathrm{vtx}}/\mathrm{ndf} > 70$ to select candidates where the two muons come from different vertices and thus resemble the heavy quark background. Figures 11.11(e) and (f) show the invariant $m_{\mu^+\mu^-}$ distribution for those samples, normalised to the number of $Z \to \mu^+\mu^-$ candidates (*cf.* Figs 11.11(a) and (b)) in the signal mass range between 60 and 120 GeV/c^2. The fraction of background candidates is again estimated from a fit by an exponential function integrated over the signal range. The resulting values are $(0.10 \pm 0.05) \times 10^{-3}$ for pA and $(0.18 \pm 0.05) \times 10^{-3}$ for Ap. The efficiency of the χ^2_{vtx}-cut for heavy quark background is estimated from $b\bar{b} \to \mu^+\mu^- X$ MC to be (9.6 ± 3.7) %. So the final fractions obtained are $(1.04 \pm 0.63) \times 10^{-3}$ for pA and $(1.88 \pm 0.52) \times 10^{-3}$ for Ap.
The second method selects heavy quark candidates by using an anti-isolation cut on the muon candidates. The isolation is defined as the fraction $z = p_{\mathrm{T}}(\mu)/p_{\mathrm{T}}(\mathrm{jet})$ of transverse momentum carried by the muon in a jet built by tracks around the muon candidate. The jet is reconstructed with the anti-k_t algorithm [101] with a cone size of $R = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2} = 0.5$ where $\Delta\phi_{ij}$ and $\Delta\eta_{ij}$ are the separation of two tracks in the jet in azimuthal angle and pseudorapidity, respectively. Well-isolated muons have values of z close to 1.
The anti-isolation is applied by requiring that both muon candidates have a value of $z < 0.7$. Figures 11.11(g) and (h) show the resulting invariant mass distributions, reweighted for the track-multiplicity and normalised to the total number of $Z \to \mu^+\mu^-$ candidates (*cf.* Figs 11.11(a) and (b)) in the signal mass range between 60 and 120 GeV/c^2. The fraction of background candidates is also in this case determined by the integration over the signal region of an exponential function that has been fit to the invariant mass distribution. The obtained fractions are $(0.32 \pm 0.08) \times 10^{-3}$ for pA and $(0.41 \pm 0.05) \times 10^{-3}$ for Ap. This corresponds to fractions after the efficiency correction for the anti-isolation cut of (23.5 ± 2.3) % – also estimated from $b\bar{b} \to \mu^+\mu^- X$ MC – of $(1.36 \pm 0.34) \times 10^{-3}$ for pA and $(1.74 \pm 0.21) \times 10^{-3}$ for Ap.

The results of the two methods are in good agreement. The weighted averages of the two methods for heavy quark background are added to the value estimated for the background from mis-identified hadrons leading to overall values for the purity of 0.9974 ± 0.0006 for pA and 0.9963 ± 0.0005 for Ap. The quoted uncertainty on the determined purity are the propagated uncertainties on the individual fractions of the two background components (mis-id and heavy quark), which are treated uncorrelated due to the individual data samples for the estimation. For both components the uncertainties reflect the uncertainty on the fit parameter of the exponential distribution. In the case of the heavy quark background an additional part of the uncertainty comes from the uncertainty of the efficiency corrections due to the anti-isolation cut or the cut on the vertex fit quality. This second part is added in quadrature to the part coming from the uncertainty from the exponential fit. Further also a description of the invariant mass distribution with the sum of two exponential distributions is tried for the samples describing the heavy quark background showing an identical result of the final fraction of background candidates.

There is a small fraction of $b\bar{b}$ candidates that appear as same-sign candidates (*e.g.* by B^0/B_s^0-oscillation or cascading decays of one b quark, *i.e.* $b/\bar{b} \to \bar{c}/cX \to \mu^-/\mu^+X$). These candidates are included in the purity estimation using the same-sign candidates although they do not appear as background in the sample of Z candidates. Figure 11.13 shows the invariant dimuon mass distributions for opposite- and same-sign candidates in $b\bar{b}$ MC. From these distributions it can be concluded that this inclusion has a negligible effect on the purity.

It is not obvious that the purity for collisions at $\sqrt{s} = 5\,\text{TeV}$ can be estimated from collisions at $\sqrt{s} = 8\,\text{TeV}$ as the kinematics are different and the cross sections of the background processes do not scale exactly as the signal process as a function of the centre-of-mass energy.
As a first cross check the amount of background attributed to mis-identification is estimated from the sample of Z candidates where at least one of the muon candidates is not identified as a muon. To model the mis-identification probability, the not identified muon candidates are weighted according to the function

$$w(p) = (1 - \exp(-p_0/p)), \tag{11.2}$$

where p is the momentum of the muon candidate, and p_0 is η-dependent parameter listed in Tab. 11.3. The numbers of p_0 have been determined from minimal bias data in pp collisions [20]. The resulting invariant dimuon mass distributions in the range from 40 to $160\,\text{GeV}/c^2$ for Ap is shown in Fig. 11.14(a). An exponential function fitted to the data is used to estimated the background fraction based on mis-identification. For pA it is $(1.9 \pm 0.8) \times 10^{-3}$ and for Ap it is $(4.2 \pm 0.7) \times 10^{-3}$. These numbers are in good agreement with the predictions based on the pp collisions at $\sqrt{s} = 8\,\text{TeV}$. The quoted uncertainties are only statistical.
As a second cross check the invariant mass distribution of same-sign dimuon candidates in proton-lead data has been studied. Due to the limited statistics of candidates when applying the full selection (two candidates in Ap and zero candidate in pA) the fraction of background candidates is estimated from the samples without the cuts on E/pc, IP_{PV} and $\text{Prob}(\chi^2_{\text{trk}},\text{ndf})$. The corresponding invariant dimuon mass distribution for Ap is shown in Fig. 11.14(b). The efficiency of these cuts for same-sign candidates, evaluated in pp data, is (0.83 ± 0.23) %. This leads to an expected impurities of $(2.3 \pm 0.9) \times 10^{-3}$ in pA and $(6.1 \pm 1.2) \times 10^{-3}$ in Ap, where the quoted uncertainties are statistical only. Although the statistics is very small, it can be concluded that no significant enhancement of background events with respect to the predictions from the pp

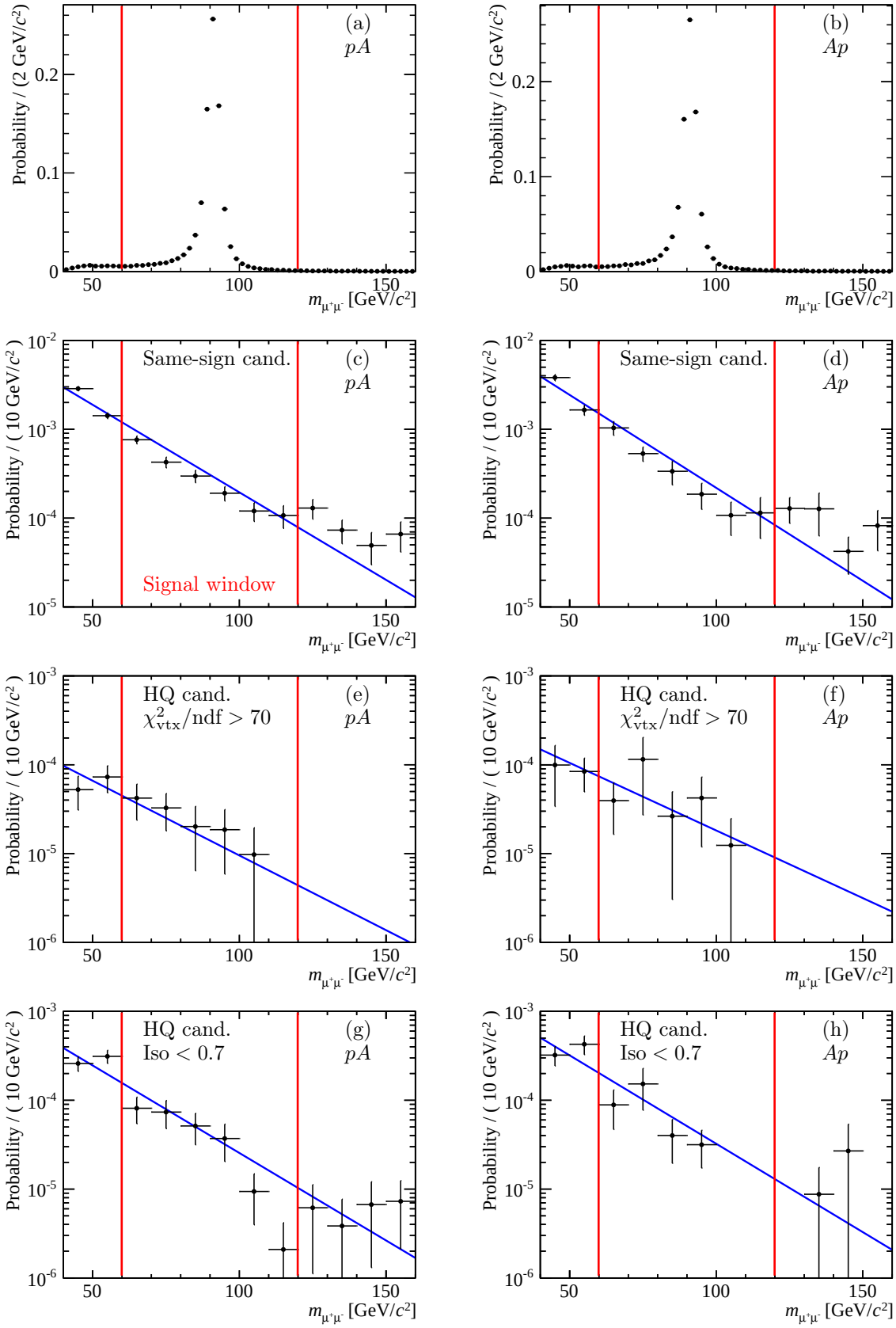


Figure 11.11: The invariant dimuon mass distributions used to estimate the background contamination reweighted for the track multiplicity in (left column) pA and (right column) Ap. The first row shows the invariant mass distribution for $Z \to \mu^+\mu^-$ candidates in pp collisions. The second row shows the invariant mass distribution for same-sign candidates in pp collisions while the third row shows the distributions for candidates in pp collisions with $\chi^2_{\rm vtx}/{\rm ndf} > 70$. The last row shows the invariant mass distribution for candidates with a track isolation of the two muons smaller than 0.7 in pp collisions. The blue curves in Figs (c) to (h) show the fitted exponential distribution used to estimated the fraction of background events in the signal window.

collisions at $\sqrt{s} = 8\,\text{TeV}$ is observed.
Furthermore, the number of candidates not passing the cuts of E/pc, IP_{PV} and $\text{Prob}(\chi^2_{\text{trk}},\text{ndf})$ in the region $40 < m_{\mu^+\mu^-} < 120\,\text{GeV}/c^2$ is in good agreement with the expected number of combinatorial background combinations based on the difference in track multiplicity scaling with n^2_{tracks}.
Although all this cross checks did not give a hint for any deviation of the purity from the values determined with pp data, a systematic uncertainty is assigned to the usage of pp collision data at $\sqrt{s} = 8\,\text{TeV}$. It is estimated as the expected change in the purity if the cross sections of all the relevant background process is the same for $\sqrt{s} = 5\,\text{TeV}$ as of $\sqrt{s} = 8\,\text{TeV}$ although these cross sections are increasing similarly as the Z production cross section as a function of $\sqrt{s}$ (*cf.* Fig. 1.3). The resulting systematic uncertainty is for both beam configurations 0.005, which is one order of magnitude larger than the values determined above based on the statistical uncertainty due to the size of the data sample used to extract the purity from pp data. So the final values for the purity are 0.997 ± 0.005 for pA and 0.996 ± 0.005 for Ap.

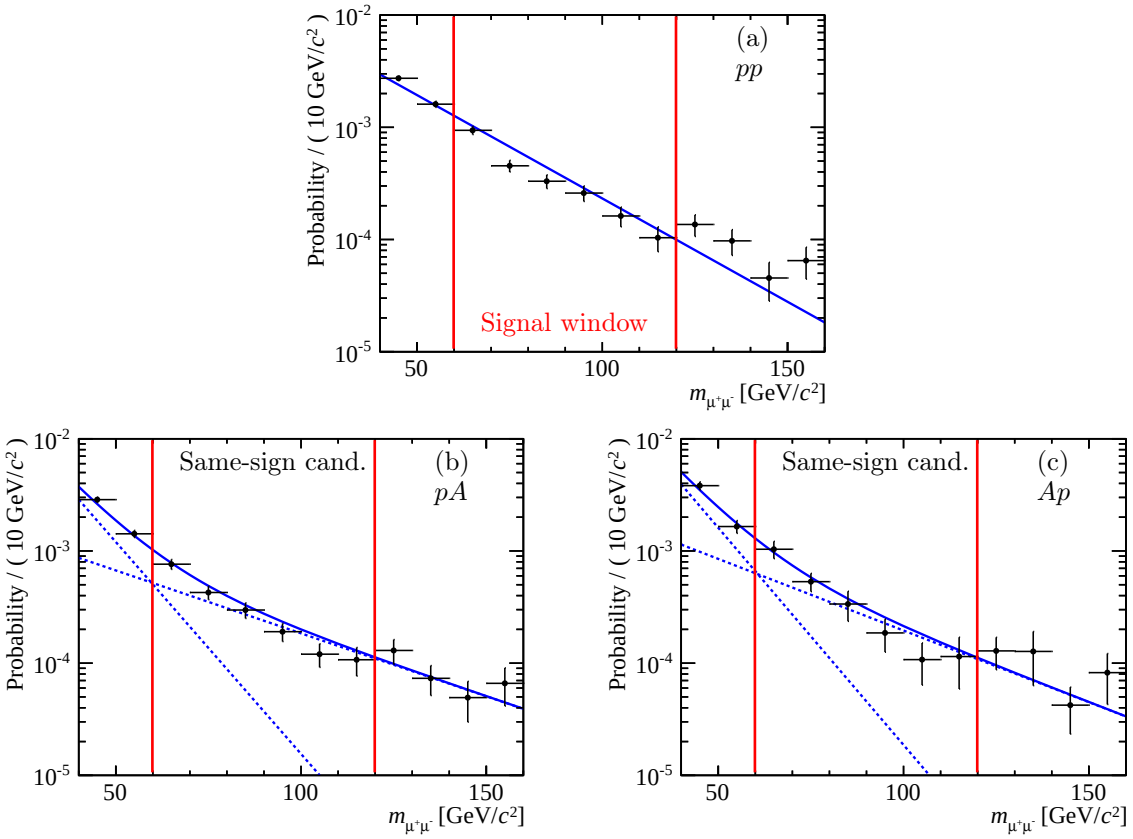


Figure 11.12: (a) The invariant dimuon mass distribution for same-sign candidates in pp collisions without multiplicity reweighting. Figures (b) and (c) show the invariant dimuon mass distributions in pp collisions reweighted for the track multiplicity in pA and Ap, respectively. The solid blue line shows the fitted sum of two exponential distributions (the individual components are shown by the dashed lines) to quantify the impact of the non-exponential behaviour of the data at high $m_{\mu^+\mu^-}$ values using a single exponential function.

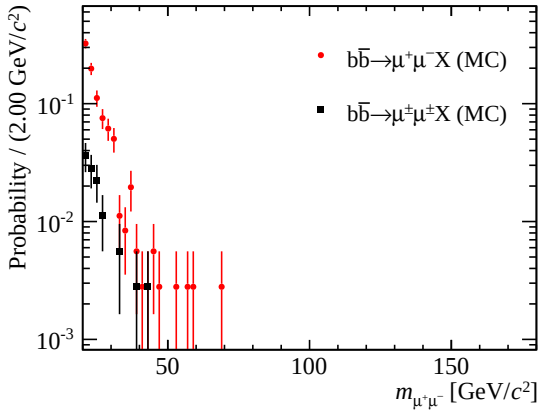


Figure 11.13: Distribution of opposite- and same-sign dimuon candidates in $b\bar{b}$ MC normalised to the total number of dimuon candidates.

Table 11.3: Muon mis-identification probability parameter. (taken from Ref. [20])

Muon pseudorapidity	p_0 **[GeV/**c**]**
$[2.0, 2.5]$	0.168 ± 0.128
$[2.5, 3.0]$	0.258 ± 0.045
$[3.0, 3.5]$	0.244 ± 0.019
$[3.5, 4.0]$	0.242 ± 0.011
$[4.0, 4.5]$	0.275 ± 0.008

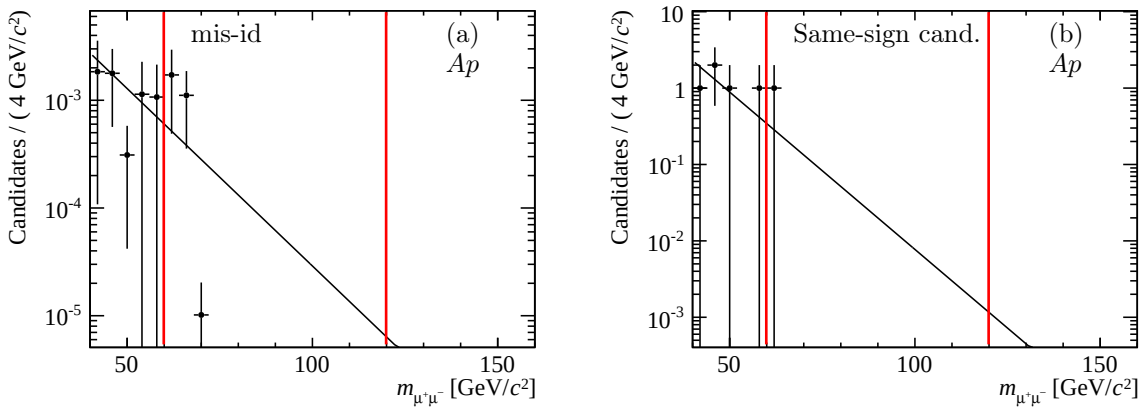


Figure 11.14: Invariant dimuon mass distribution (a) for candidates with at least one mis-identified track reweighted for the mis-identification probability as described in the text and (b) for same-sign candidates with no additional cuts on E/pc, $\mathrm{IP}_{\mathrm{PV}}$ and $\mathrm{Prob}(\chi^2_{\mathrm{trk}},\mathrm{ndf})$ in Ap fitted with an exponential function. The red lines show the signal window.

11.4 Efficiency Determination

The GEC efficiency is determined from the samples of $J/\psi \to \mu^+\mu^-$ candidates while all other efficiencies are determined by tag-and-probe methods using the sample of $Z \to \mu^+\mu^-$ candidates in pp data at $\sqrt{s} = 8\,\text{TeV}$ with multiplicity reweighting.

11.4.1 Global Event Cut Efficiency

The relevant GECs in the pA and Ap samples are the number of OT clusters (smaller than 15'000) and the number of IT clusters (smaller than 3000). The cut on the number of VELO clusters (smaller than 6000) does not need to be considered as it is 100 % efficient for both beam configurations.
The estimation of the GEC efficiency is done with a sample of background subtracted (${}_s\mathcal{P}lot$ technique) $J/\psi \to \mu^+\mu^-$ candidates in pA or Ap data. Figure 11.15 shows the distributions of the number of IT and OT clusters in the two samples. This figure further shows the corresponding distributions for $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ candidates in pp collisions. Differences in the distributions of the number of OT and IT clusters between $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ candidates, which appear due to the different production mechanisms, are taken into account by reweighting the distributions from $J/\psi \to \mu^+\mu^-$ candidates in pA/Ap data by the ratio between the corresponding distributions of $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ in pp collisions.

The resulting distributions are then fit by an empirical model based on the sum of two generalised Gamma distributions ${}_d\mathcal{GM}$.

$$\begin{aligned} {}_d\mathcal{GM}(x|f,\beta_1,\beta_2,\gamma_1,\gamma_2,\mu_1,\mu_2) &= f\cdot\mathcal{GM}(x|\beta_1,\gamma_1,\mu_1) + (1-f)\cdot\mathcal{GM}(x|\beta_2,\gamma_2,\mu_2) \\ \text{with} \qquad \mathcal{GM}(x|\beta,\gamma,\mu) &= \frac{\left(\frac{x-\mu}{\beta}\right)^{\gamma-1}\exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\Gamma(\gamma)}, \end{aligned} \tag{11.3}$$

where Γ is the Gamma function and $\gamma, \beta > 0$, $x > \mu$ (*cf.* Fig. 11.16). The usage of Gamma distributions to describe these distributions has been already tested and applied in previous analysis like the measurement of Z production in pp collisions [190] or the measurement of J/ψ production in pPb collisions [35]. From the fitted function the fractions of events above the GEC thresholds are determined shown in Tab. 11.4. Table 11.4 also shows the correlation ρ between the number of IT and OT clusters. The overall efficiency is estimated by

Table 11.4: Summary of the Global Event Cut efficiencies. $\varepsilon^{\text{GEC,IT/OT}}$ are the individual efficiencies for the cuts on the number of IT and OT clusters while ε^{GEC} is the combined efficiency. ρ is the correlation between the number of IT and OT clusters.

Configuration	$\varepsilon^{\text{GEC,OT}}$	$\varepsilon^{\text{GEC,IT}}$	ρ	ε^{GEC}
pA	1.000	1.000	0.849	1.000
Ap	0.976	0.994	0.958	0.978

$$\varepsilon^{\text{GEC}} = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \mathrm{d}x_1'\mathrm{d}x_2'\mathcal{N}(x_1', x_2'|\rho), \tag{11.4}$$

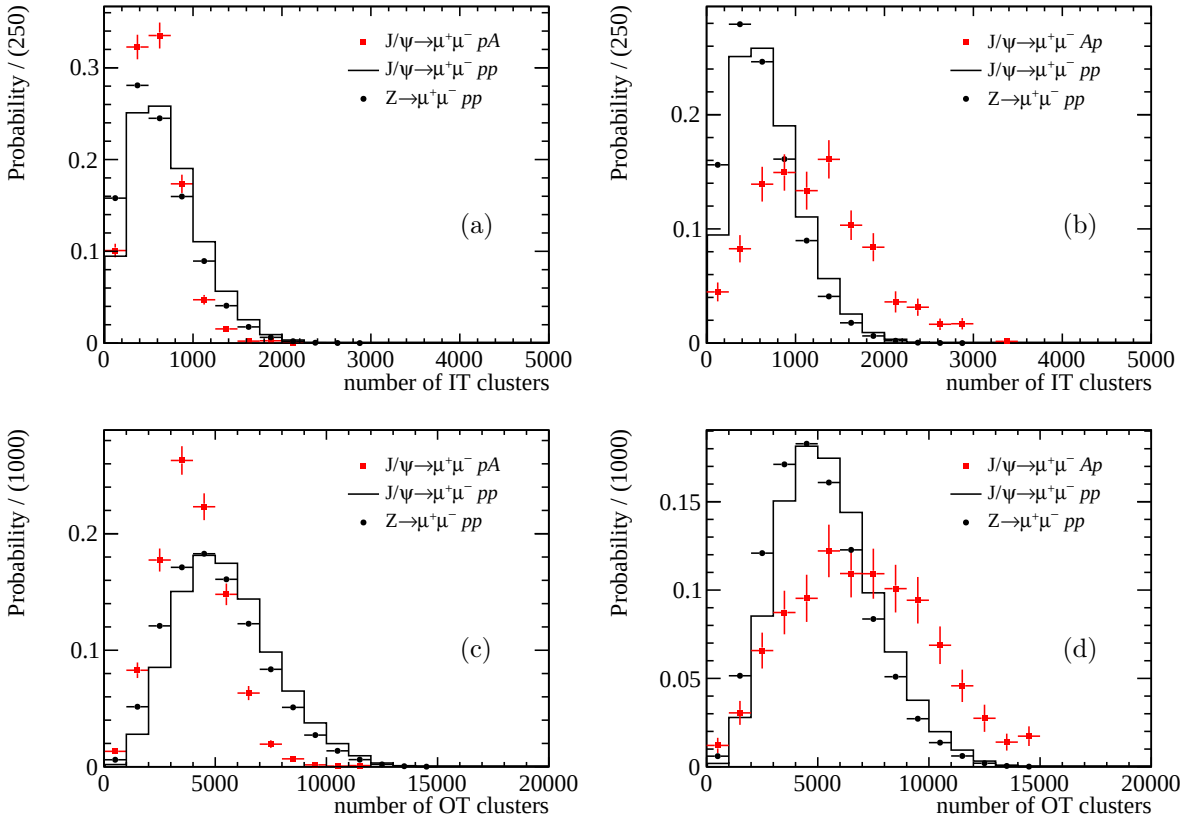


Figure 11.15: The top row shows the distributions of the number IT clusters for (a) pA and (b) Ap data while the bottom row shows the distributions of the number of OT clusters for (c) pA and (d) Ap data. The data points are shown in red while the black points show the corresponding distributions for $Z \to \mu^+\mu^-$ in pp collisions, the solid line the ones for $J/\psi \to \mu^+\mu^-$ in pp collisions.

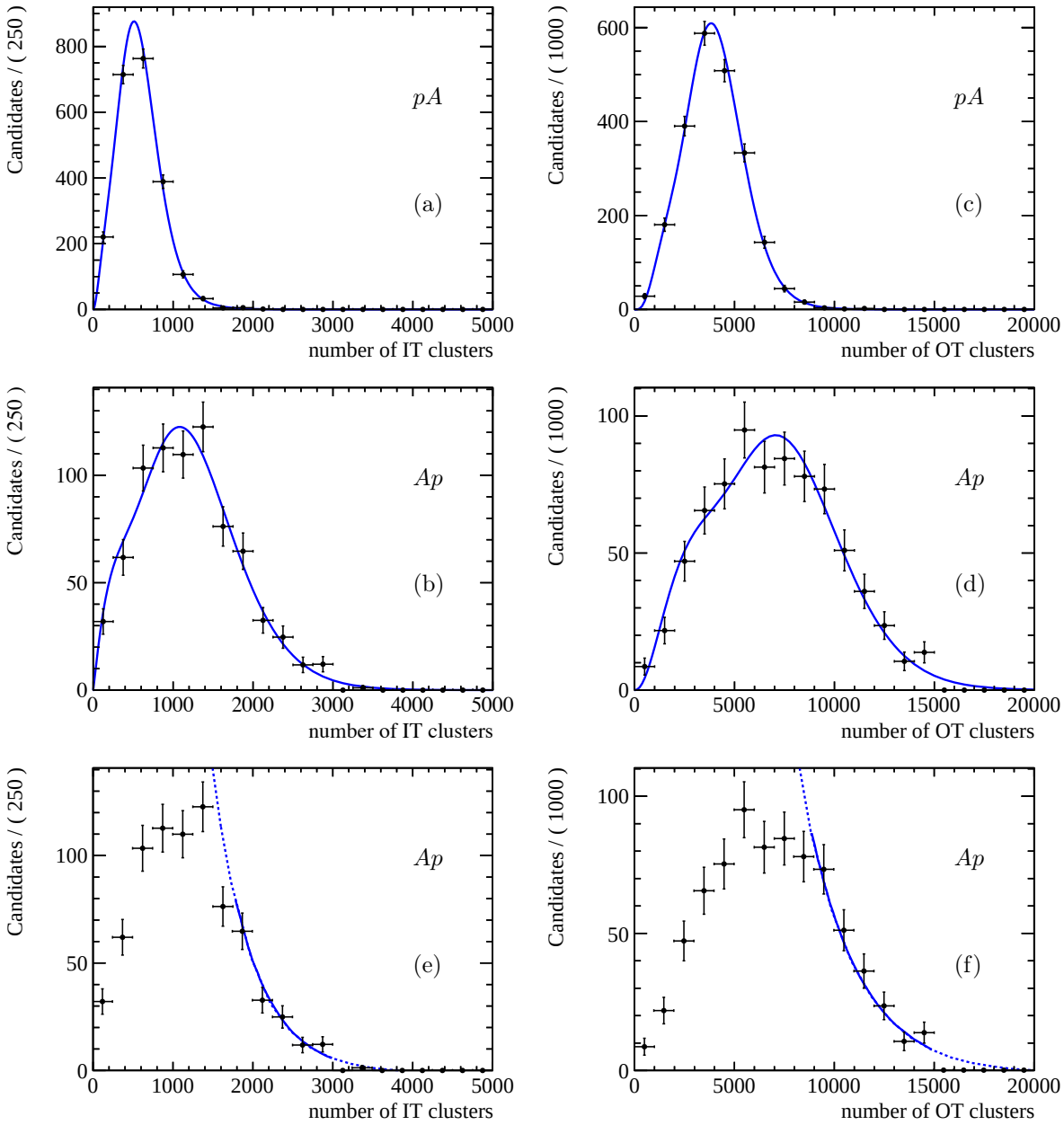


Figure 11.16: The top row shows the background subtracted distributions of (a) IT and (b) OT clusters from J/ψ candidates in pA. The distributions are reweighted for the track multiplicity difference between Z and J/ψ production in pp collisions and are fit by the sum of two Gamma distributions. The center row shows the same for (c) IT and (d) OT clusters in Ap while the bottom row ((e) and (f)) shows the same distributions as in (c) and (d), but using an exponential distribution to estimate the fraction of events removed by the GEC.

where $\mathcal{N}$ is a two-dimensional normal distribution and $x_1 = F^{-1}(\varepsilon^{\mathrm{GEC,IT}})$ and $x_2 = F^{-1}(\varepsilon^{\mathrm{GEC,OT}})$. F is the one-dimensional cumulative Normal distribution and $\varepsilon^{\mathrm{GEC,OT}}$ and $\varepsilon^{\mathrm{GEC,IT}}$ are the single efficiencies of the two individual GECs. The values for $\varepsilon^{\mathrm{GEC}}$ are 1.000 for pA and 0.978 for Ap. To estimate the systematic uncertainties the distributions of the number of IT and OT clusters have been fit by an exponential distribution in the high-value region (IT: [1800,3000], OT: [9000,15'000]) (*cf.* Fig 11.16(e) and (f)). The GEC efficiencies have been evaluated by the integral of the data histogram up to 1800 in IT and 9000 in OT and the integral of the exponential distribution in the region above. The final values of the GEC are 1.000 for pA and 0.959 for Ap. The differences to the numbers determined from the double Gamma distributions (pA: 0.000, Ap: 0.019) are taken as systematic uncertainties. As a further systematic cross check the efficiency is estimated by repeating the estimation without the reweighting of $J/\psi \to \mu^+\mu^-$ candidates. The difference to the values obtained with the weighting is zero for pA and 0.012 for Ap. As this effect is dominated by the one associated with the form of the distribution, it is not added to the systematic uncertainty.

11.4.2 Tracking Efficiency $\varepsilon^{\mathbf{REC}}$

The tracking efficiency as well as the trigger, selection and muon-identification efficiencies – described in the following subsections – are estimated by tag-and-probe methods applied to the two muon candidates. The efficiency ε is determined from the sample of $Z \to \mu^+\mu^-$ candidates in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$. This can be done as the reconstruction of the events in the pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ and the one in pA/Ap collisions use the same reconstruction version (Reco14). The efficiency is defined as the fraction between the yield of candidates, N_{tag}, with requirements on the tag muon and the yield of candidates, $N_{\mathrm{tag\&probe}}$, with requirements on the probe muon in addition to the ones in the tag sample

$$\varepsilon \;=\; \frac{N_{\mathrm{tag\&probe}}}{N_{\mathrm{tag}}}. \tag{11.5}$$

The uncertainty on the efficiency is derived from the binominal formula as

$$\sigma_\varepsilon \;=\; \sqrt{\frac{\varepsilon(1-\varepsilon)}{N_{\mathrm{tag}}}}. \tag{11.6}$$

In all cases, the efficiency is determined as a function of the pseudorapidity of the probe muon in ten equidistant bins from 2.0 to 4.5.
The distributions are as in the case of the estimation of the sample purity (*cf.* Sec. 11.3) reweighted to account for the different track multiplicities in pp, pA and Ap data.

Data Sample

A dedicated sample of $Z \to \mu^+\mu^-$ candidates is used to estimate the tracking efficiency. The Z candidates are built from a muon candidate reconstructed as a long track (serving as tag) and a second muon candidate with opposite charge sign reconstructed only from a track segment in the TT stations and the Muon System (serving as probe) [31, 126].
The selection criteria are that the p_{T} of both muon candidates are larger than $20\,\mathrm{GeV}/c$ and that

their η is between 2.0 and 4.5. It is further required that the tag muon has fired the L0Muon, the Hlt1SingleMuonHighPT and Hlt2SingleMuonHighPT trigger and that both tracks are identified as muons (*i.e.* IsMuon). Additional criteria are that the absolute difference in the azimuthal angle between the two muons $|\Delta\phi|$ is larger than 0.1, to reject candidates where the two tracks are clones, and that the summed p_{T} of tracks clustered by the anti-k_t algorithm in a cone of $R = 0.5$ around the tag muon is smaller than $2\,\mathrm{GeV}/c$.

Efficiency Determination

The yield of tag candidates N_{tag} is the number of candidates that fulfil the above selection and fall into the invariant dimuon mass window of $m_{\mu^+\mu^-} \in [70, 110]\,\mathrm{GeV}/c^2$. It has to be mentioned that due to the TTMuon track the mass resolution is significantly worse than for candidates reconstructed from two long tracks. The yield for $N_{\mathrm{tag\&probe}}$ corresponds to the number of candidates passing the same selection as the tag candidates and where the TTMuon track is matched to a long track. Further it is required that the invariant dimuon mass calculated from the track of the tag muon and the long track associated to the probe muon is larger than $40\,\mathrm{GeV}/c^2$. This additional criterion removes candidates with an accidental match between the TTMuon track and a random long track. The fraction of candidates removed by this requirement is about 0.5 %. Figure 11.17 shows as an example the invariant mass distributions of tag and probe candidates for the lowest and highest pseudorapidity bin.
From these distributions it is not obvious that all the candidates can be classified as signal and that thus there is no need to model the background. Figure 11.18 shows for different η bins the invariant mass distribution for Long-TTMuon candidates as well as the invariant mass distribution from candidates build from two long tracks (Long-Long) that are identified as signal (*cf.* Fig. 11.19), but where the invariant mass has been calculated after demoting one of the two muons to a TTMuon track. As the distributions are in agreement for $m_{\mu^+\mu^-} > 70\,\mathrm{GeV}/c^2$, a modelling of background contributions is not deemed necessary.

The tracking efficiency as a function of the muon η for positive and negative charged muons as well as for reweighted track multiplicities matching the one in pA and Ap data is shown in Fig. 11.20.
The expected number of Z candidates in the collected proton-lead data is small. Thus, the efficiency correction in the cross section calculation is not performed on an event-by-event basis, but using the integrated efficiency over the considered muon pseudorapidity range, the tracking efficiency as a function of η must be folded with the η distribution of the muons.
The η distribution of the muons is taken from MC samples generated by Pythia8 [217] with the MSTW08 LO PDF set [195]. The samples used are from simulated pp and pn collisions where the first proton has an energy of 4 TeV and the other proton or the neutron – mimicking the constituents of the lead nucleus – has an energy of 1.58 TeV. Figure 11.21 displays the resulting η distribution of both muons together with the corresponding distributions in pp collisions with $\sqrt{s} = 8\,\mathrm{TeV}$ (*i.e.* $4 + 4\,\mathrm{TeV}$) and 3.16 TeV (*i.e.* $1.58 + 1.58\,\mathrm{TeV}$).
The selections cuts applied on the muons are $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $2.0 < \eta < 4.5$ while the Z candidate must fulfil $60 < m_{\mu^+\mu^-} < 120\,\mathrm{GeV}/c^2$.
For pA collisions, the distribution is very similar to the one from pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$.

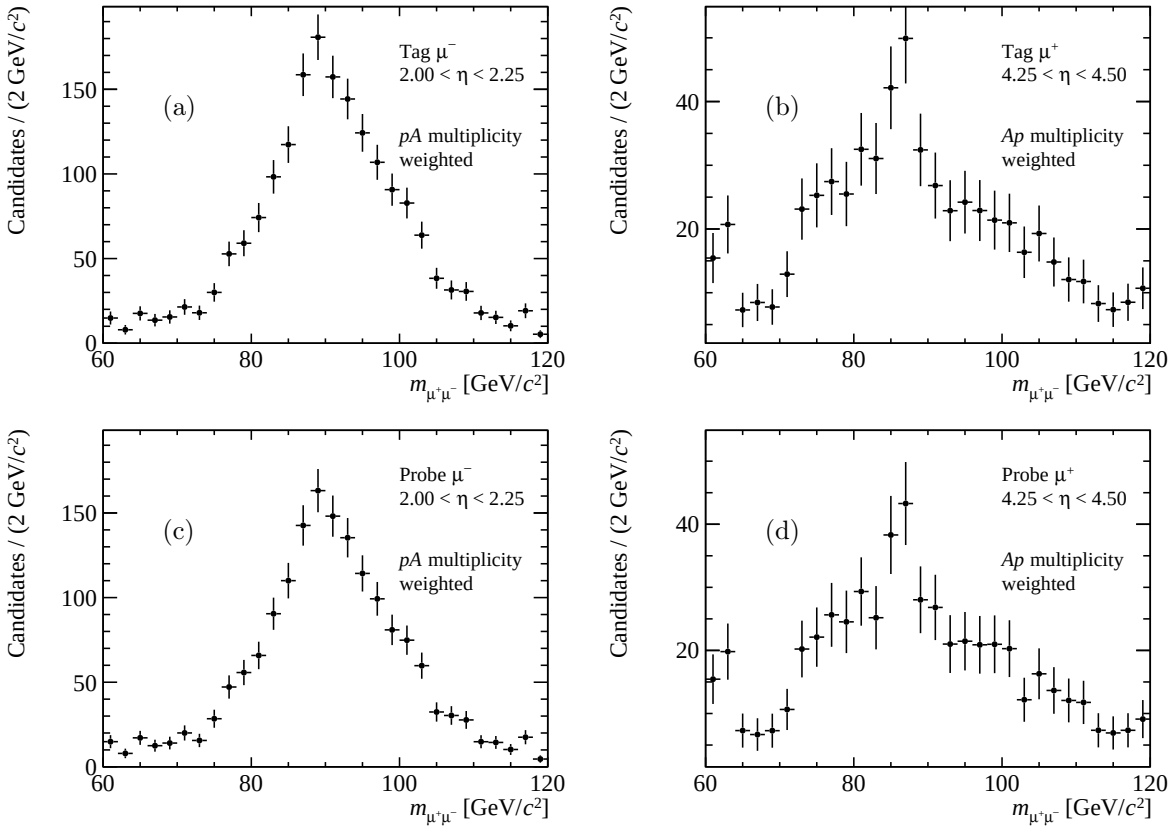


Figure 11.17: Invariant dimuon mass distribution of the (a) tag and (c) probe sample for the determination of the tracking efficiency of the negative charged muon in the lowest η bin ($2.00 < \eta < 2.25$) and for a reweighted tracking multiplicity matching the one of pA data. The tag muon is reconstructed as a long track while the probe muon is reconstructed as a TTMuon-track. Figures (b) and (d) show the invariant dimuon mass distribution for the positive charged muon as probe in the highest η bin ($4.25 < \eta < 4.50$) and for a reweighted track multiplicity comparable to the one in Ap data.

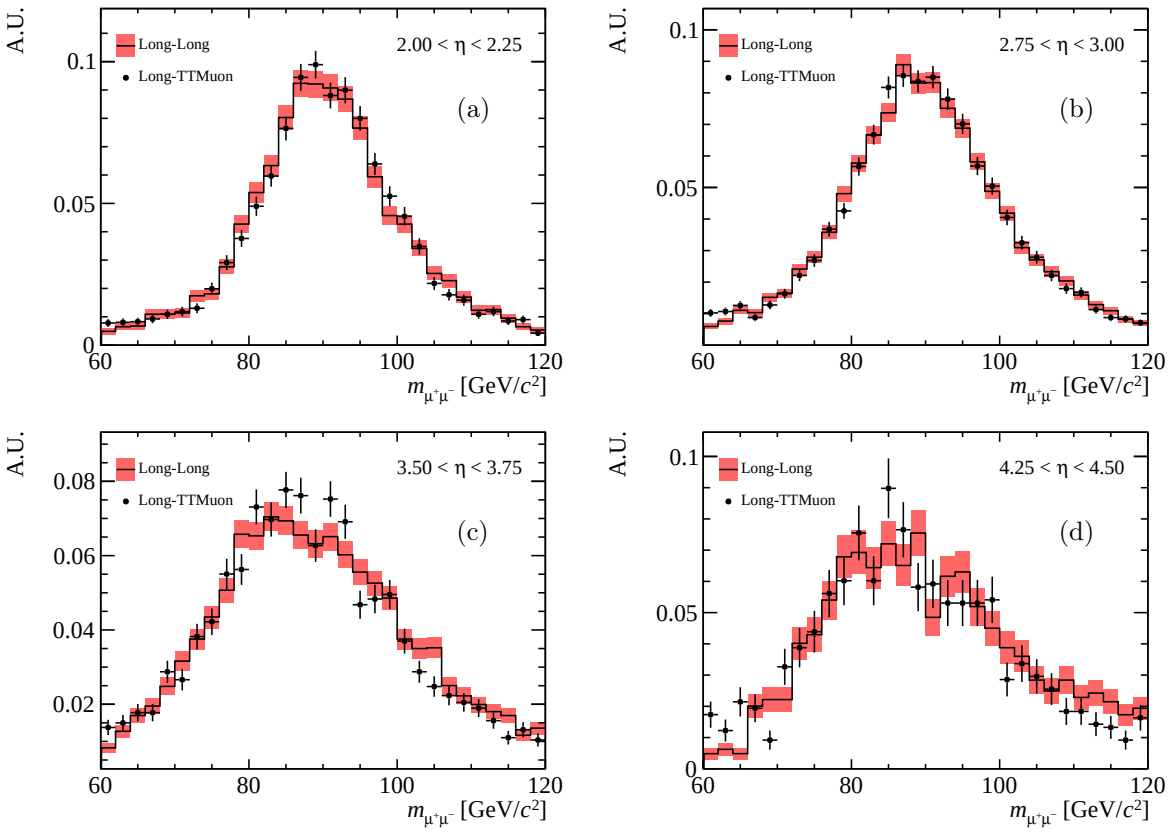


Figure 11.18: Comparison of the invariant mass distributions for Long-TTMuon combinations and Long-Long combinations with one muon demoted to a TTMuon track for different η ranges of the TTMuon track: (a) $2.00 < \eta < 2.25$, (b) $2.75 < \eta < 3.00$, (c) $3.50 < \eta < 3.75$, (d) $4.25 < \eta < 4.50$. The distributions are normalised to the region of $70 < m_{\mu^+\mu^-} < 110\,\mathrm{GeV}/c^2$, in which they show a good agreement. This means that there is no significant amount of background in the Long-TTMuon sample.

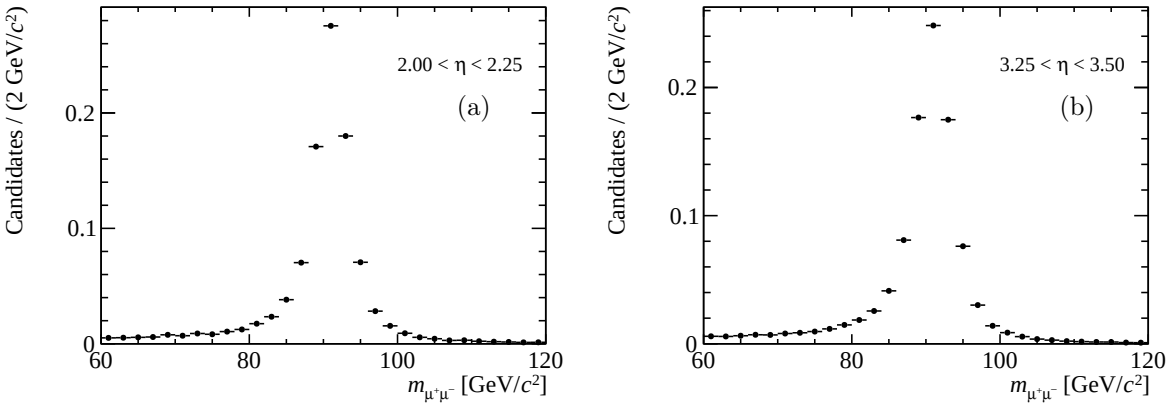


Figure 11.19: Invariant dimuon mass distribution of the Long-Long candidates used for the check of the background contamination of the sample of Long-TTMuon combinations. Figure (a) shows the distribution of candidates in which one of the muons is in the range $2.00 < \eta < 2.25$ and Fig. (b) that in which one of the muons is in the range $3.25 < \eta < 3.50$. The sample shows no significant background contribution as the candidates in the left tail are corresponding to the expectation for Drell-Yan production and for the radiative tail of the Z resonance.

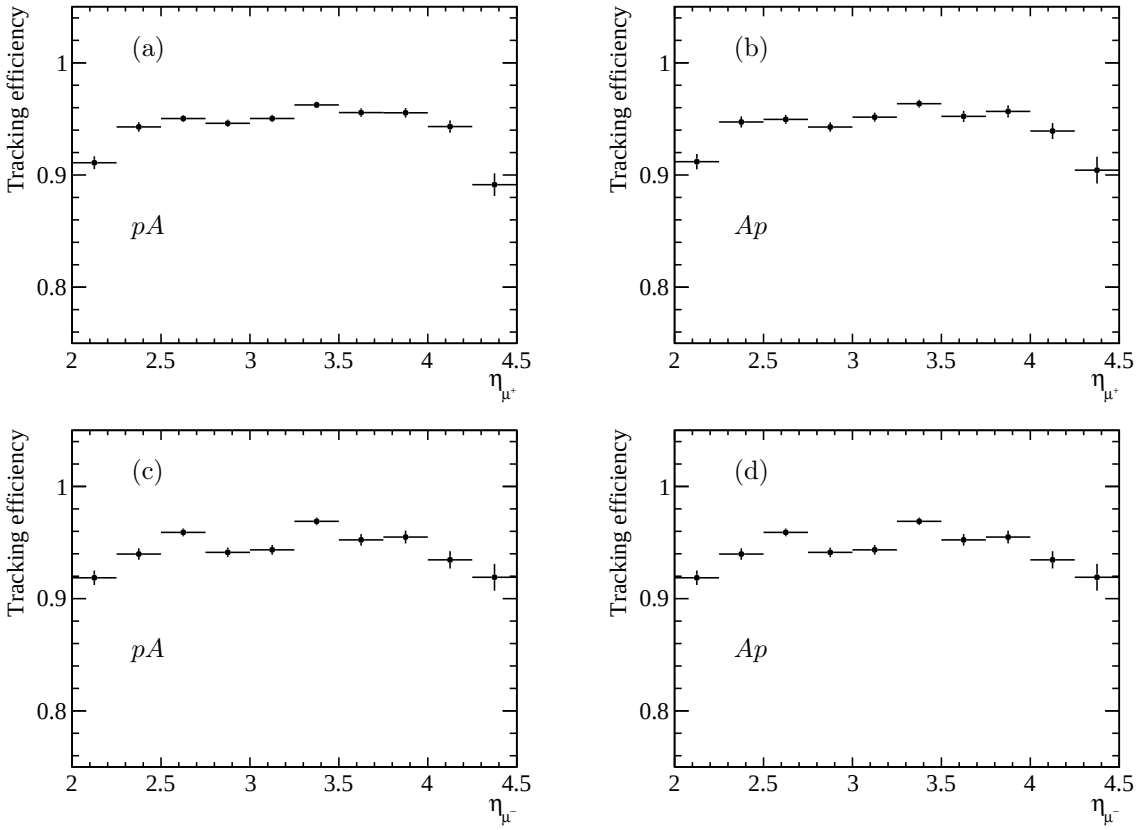


Figure 11.20: The tracking efficiency as a function of the pseudorapidity of the track for (a) positive and (c) negative charged muons from $Z \to \mu^+\mu^-$ in pA data. Figures (b) and (d) show the same for positive and negative charged muons in Ap data.

On the other hand for Ap collisions, the η distribution shows a kinematic cut-off at high η values due to the smaller energy of the nucleon and therefore a similar behaviour as for pp collisions at $\sqrt{s} = 3.16\,\text{TeV}$.
The final η distribution for proton-nuclei collisions is estimated by

$$\frac{\mathrm{d}\sigma_{pA/Ap}}{\mathrm{d}\eta} = N_Z \cdot \frac{\mathrm{d}\sigma_{pp}}{\mathrm{d}\eta} + (N_A - N_Z)\frac{\mathrm{d}\sigma_{pn}}{\mathrm{d}\eta}, \tag{11.7}$$

where N_Z is the number of protons in the lead nucleus and N_A is the total number of nucleons. This procedure has been also applied when estimating the integrated selection, muon-identification and trigger efficiencies described below.
Figure 11.22 also shows the two-dimensional distribution for the pseudorapidities of the two muons used to check if there is a significant correlation between the two variables requesting for a two-dimensional calculation of the integrated efficiency. There is no significant dependance between the two pseudorapidities – the absolute correlation is in all cases smaller than 5 % – and therefore a two-dimensional convolution is not deemed necessary.

The final values for the tracking efficiency of both tracks are 0.903±0.018 for pA and 0.892±0.019 for Ap. The dominating part of the uncertainty comes from the integration procedure. The quoted systematic uncertainty arises on the one hand from the binomial uncertainty on the determined efficiency in each η bin due to the statistics of the Z candidates in the used sample. On the other hand the uncertainty on the η spectrum has been propagated to the final efficiency value. This uncertainty comes from the statistics of the MC sample used to describe the spectrum and from the uncertainty due to the choice of the PDF. The latter is estimated by changing the used PDF set from the standard MSTW08 one to the CTEQ6.6 set [200] and adding the difference in quadrature to the uncertainty from the statistics. The uncertainty on the η spectrum is later also used in the estimation of the other efficiencies.
As cross checks the impact of a change of the $|\Delta\phi|$ cut from 0.1 to 0.2 in the selection requirements for the candidates to estimate the tracking efficiency as well as of a tighter mass window $m_{\mu^+\mu^-} \in [80, 100]\,\text{GeV}/c^2$ of the candidates has been studied. Both checks did not show any significant deviation from the determined final values of the tracking efficiency.

11.4.3 Selection Efficiency $\varepsilon^{\mathbf{SEL|REC}}$

The tag sample to evaluate the selection efficiency is based on $Z \to \mu^+\mu^-$ candidates in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ after track multiplicity reweighting. Both muons need to have a p_T larger than $20\,\text{GeV}/c$ and a η between 2.0 and 4.5. In addition, the tag muon must fulfil the selection criteria in terms of IP_PV, $\text{Prob}(\chi^2_\text{trk}, \text{ndf})$ and E/pc.
For the probe sample the requirements on the IP_PV, $\text{Prob}(\chi^2_\text{trk}, \text{ndf})$ and E/pc are also applied for the probe muon. Figure 11.23 shows the invariant mass distributions for the tag and probe samples. The yields are determined as the number of candidates in the invariant dimuon mass window of $m_{\mu^+\mu^-} \in [70, 110]\,\text{GeV}/c^2$.
The selection efficiency as a function of the η of the probe muon is shown in Fig. 11.24. The integrated numbers for the selection efficiency of both muons are 0.855±0.019 in pA and 0.853±0.022 in Ap using the η spectrum from simulation as described above.
The quoted systematic uncertainty is based on the binomial uncertainty on the efficiencies in

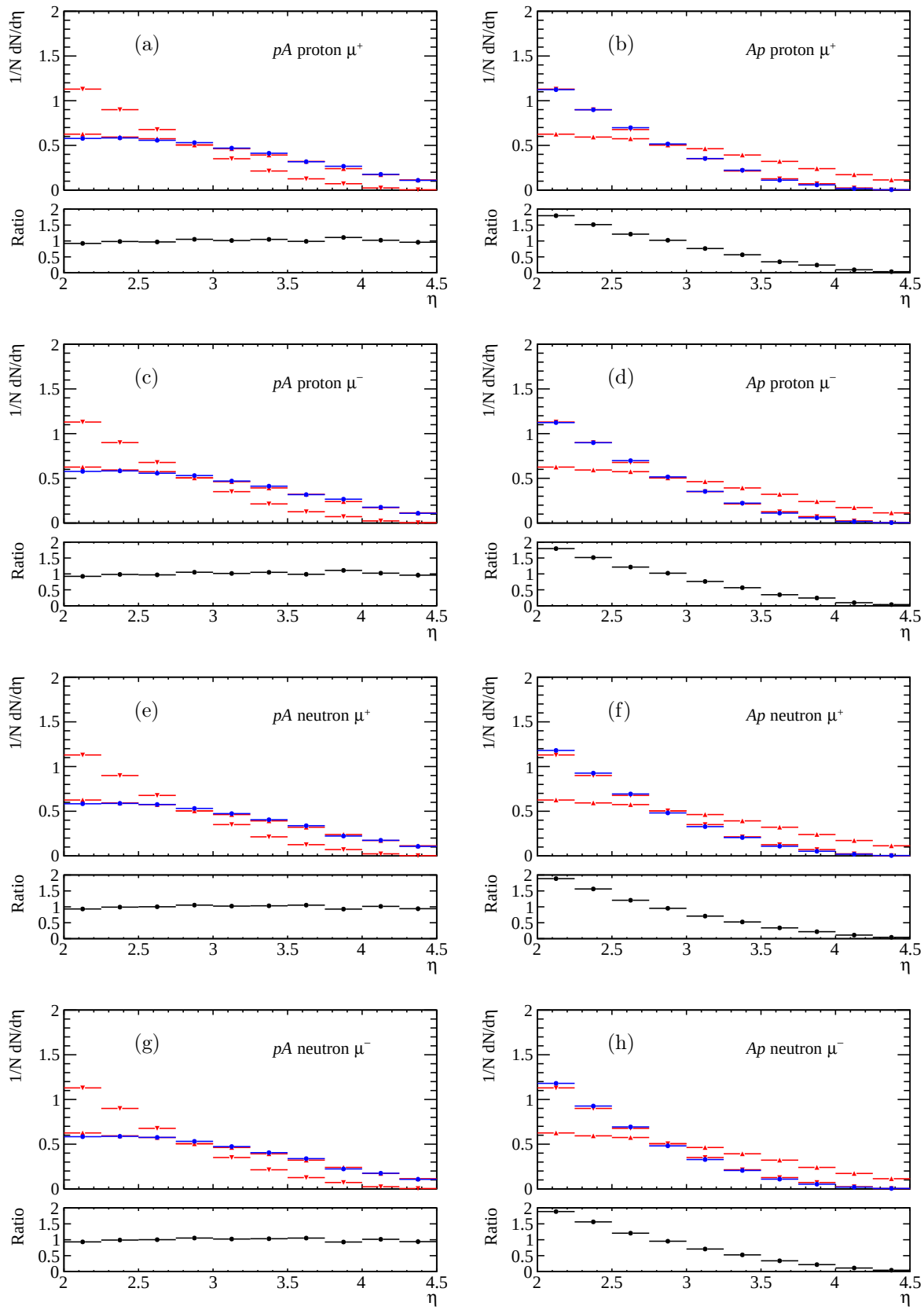


Figure 11.21: Normalised η distribution for muons (positive: (a,b,e,f), negative: (c,d,g,h)) in $Z \to \mu^+\mu^-$ simulation. The upward pointing triangles show the distribution for pp at $\sqrt{s} = 8\,\text{TeV}$, the downward one for $\sqrt{s} = 3.16\,\text{TeV}$. The dots show the distribution for simulated proton-proton (a,b,c,d) or proton-neutron collisions (e,f,g,h) at $\sqrt{s} = 5\,\text{TeV}$ ($E_p = 4\,\text{TeV}$, $E_{p,n} = 1.58\,\text{TeV}$) mimicking proton-lead collisions. The blue dots show in the right four plots the distribution in the direction of the lower energetic beam (*i.e.* *Ap* configuration), in the left four in the direction of the higher energetic beam (*i.e.* *pA* configuration). The small plots show the ratio of the distribution mimicking proton-nucleus collisions to that for pp at $\sqrt{s} = 8\,\text{TeV}$.

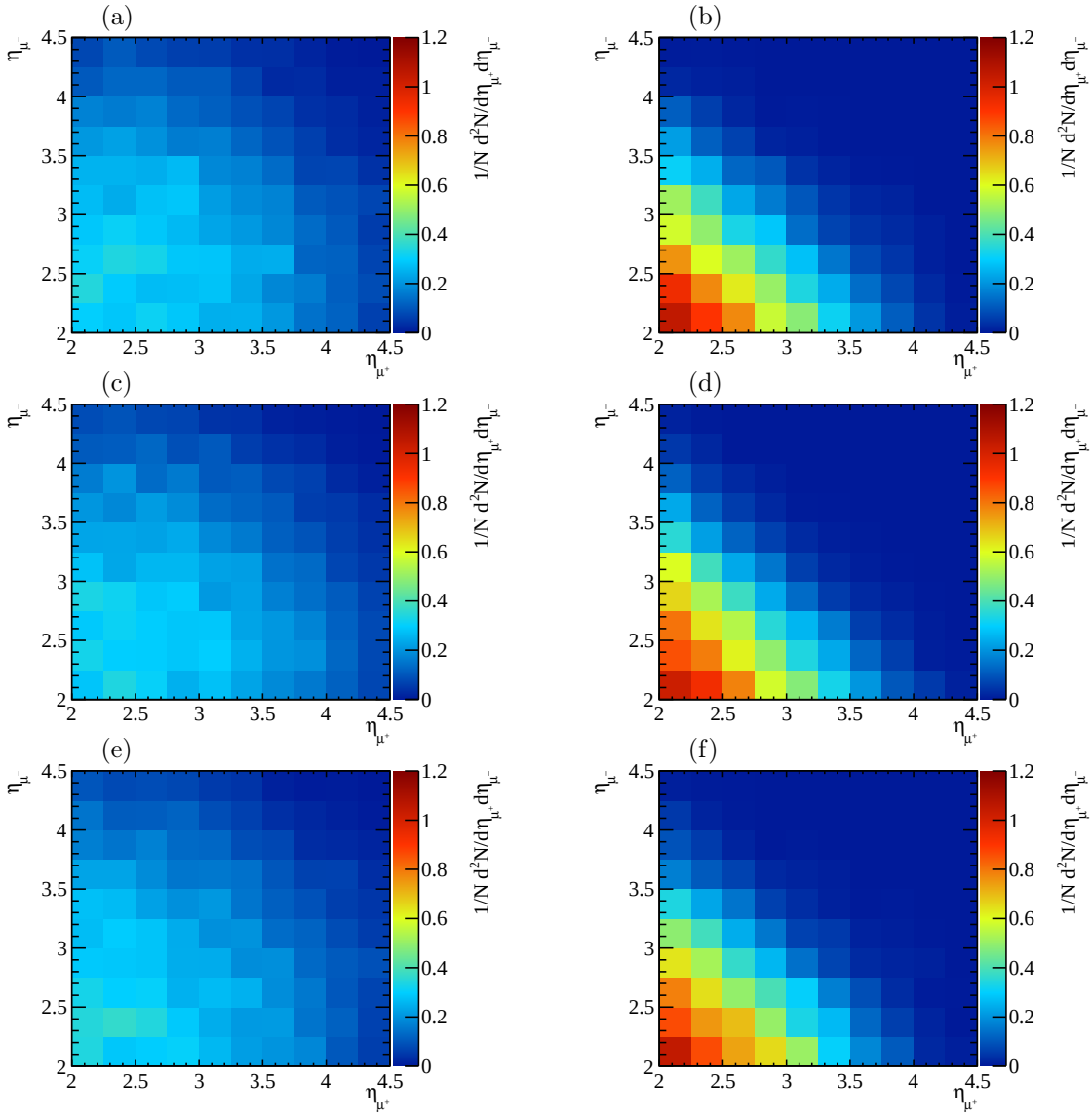


Figure 11.22: Two-dimensional normalised η distribution of the two muons in $Z \to \mu^+\mu^-$ from simulation: (a) pp at $\sqrt{s} = 8\,\text{TeV}$, (b) pp at $\sqrt{s} = 3.16\,\text{TeV}$, (c) pp at $\sqrt{s} = 5\,\text{TeV}$ (pA configuration), (d) pp at $\sqrt{s} = 5\,\text{TeV}$ (Ap configuration), (e) pn at $\sqrt{s} = 5\,\text{TeV}$ (pA configuration), (f) pn at $\sqrt{s} = 5\,\text{TeV}$ (Ap configuration).

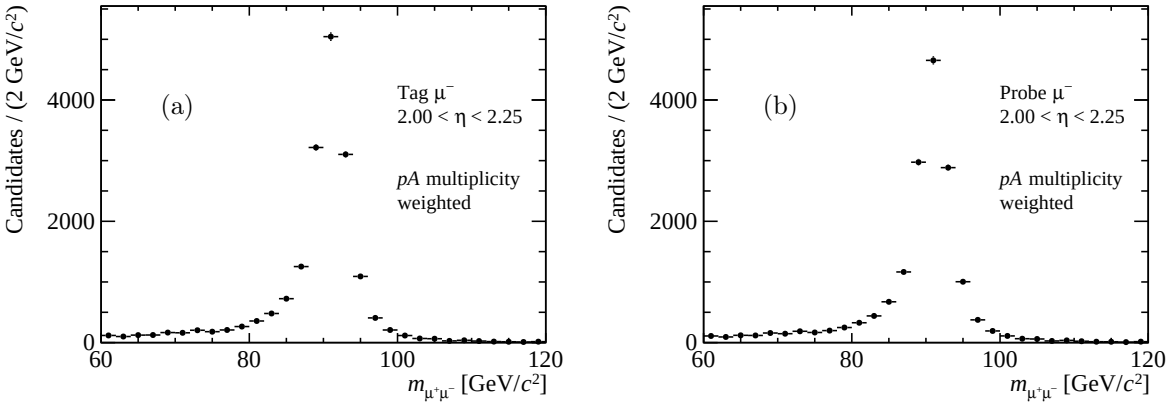


Figure 11.23: Invariant dimuon mass distribution in the (a) tag and (b) probe sample for the determination of the selection efficiency shown as example for negative charged muons with $2.00 < \eta < 2.25$ for the probe muon and multiplicity reweighting for pA.

each η bin and the uncertainty on the η spectrum determined in the same way as described in the tracking efficiency section. The two uncertainty components are propagated to the final value in the integration over the considered η range.
As a cross check the selection efficiency has been estimated with a wider ($m_{\mu^+\mu^-} \in [60, 120]\,\text{GeV}/c^2$) and tighter ($m_{\mu^+\mu^-} \in [80, 100]\,\text{GeV}/c^2$) invariant dimuon mass window, but the integrated efficiency value has not shown any significant deviation. The yields in the tag and in the probe samples are estimated as a further cross check from fits of the invariant dimuon mass distributions (*cf.* Fig. 11.25). The fit model for the resonant component of the signal is the convolution of the sum of two Crystal Ball functions with common central value and width and a Breit-Wigner distribution with a mean of zero and a width Γ fixed to the nominal Z width of $2.495\,\text{GeV}/c^2$ [81]. The tails of the Crystal Ball functions are to the opposite sides of the mass distribution and their parameters are fixed to the values determined from simulation (*cf.* Sec. 8.2.1). The background is parametrised by an exponential distribution including also the non-resonant Drell-Yan component. The efficiency evaluated from the yields determined by the fit is within the uncertainties in agreement with the values quoted above.
The same cross check has been also performed for the muon-identification and trigger efficiency below, showing also agreement with the values determined from the method based on the counted yields.

11.4.4 Trigger Efficiency $\varepsilon^{\text{TRIG|SEL}}$

The trigger efficiency is evaluated on a sample of $Z \to \mu^+\mu^-$ candidates from pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with track multiplicity reweighting for pA and Ap data. The selection is the usual one described in Tab. 11.2. Further it is required in the tag sample that the tag muon has fired the L0Muon, Hlt1SingleMuonHighPT and Hlt2SingleMuonHighPT trigger. For the probe sample, the additional requirement imposed on the probe muon is that it has fired the Hlt1SingleMuonHighPT trigger[6].

[6] N.B. that the L0 and HLT2 trigger levels are in pass-through mode in proton-lead collisions

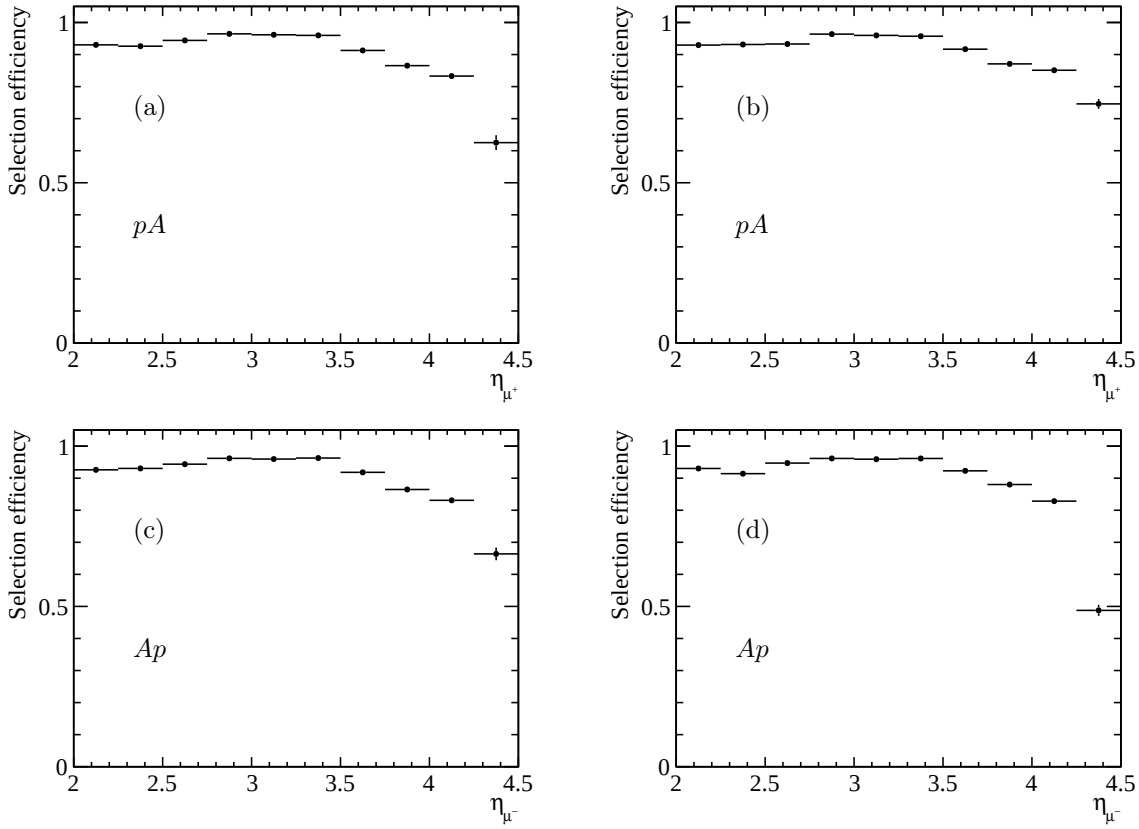


Figure 11.24: The estimated selection efficiency as a function of the pseudorapidity of the track for (a) positive and (c) negative charged muons from $Z \to \mu^+\mu^-$ in pA data. Figures (b) and (d) show the same for positive and negative charged muons in Ap data.

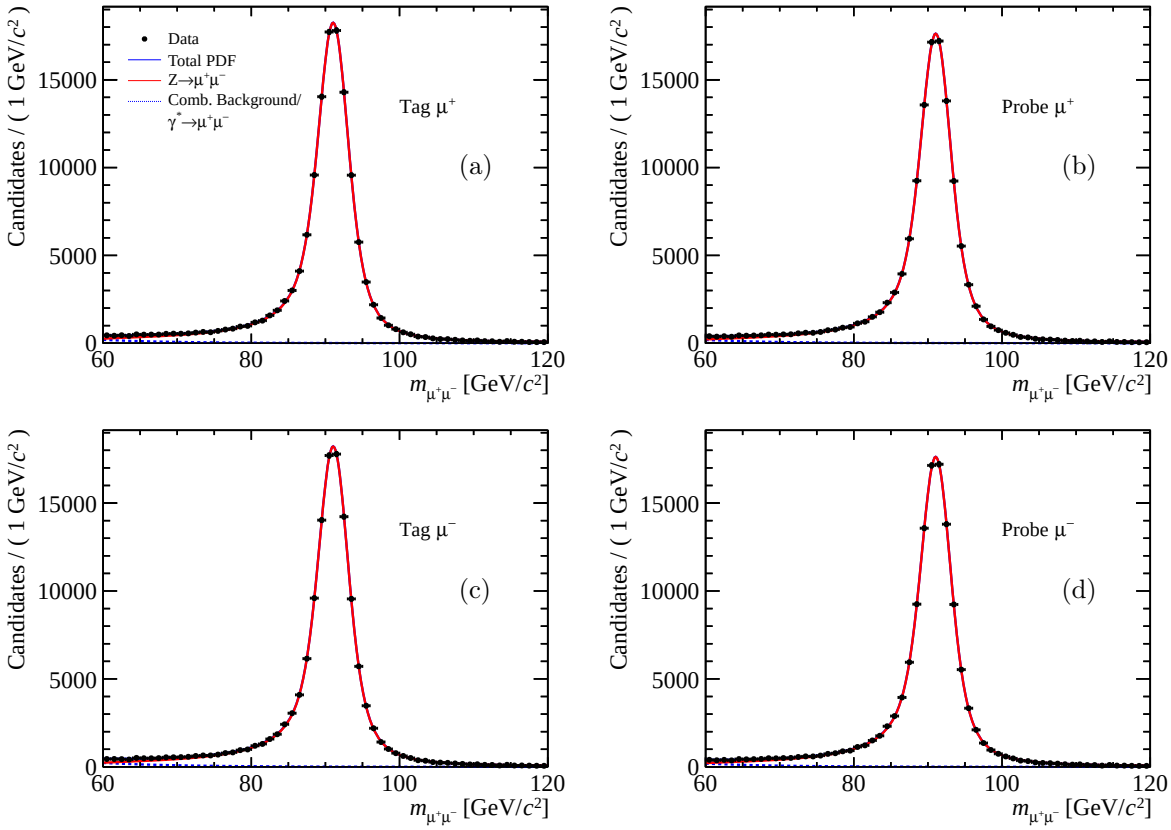


Figure 11.25: The invariant dimuon mass distribution of the tag and probe samples for both muon charges in the full pseudorapidity range for the determination of the selection efficiency. The distributions are fit with the convolution of a Breit-Wigner function and the sum of two Crystal Ball functions for signal (red solid line) and an exponential function for background (blue dashed line). The full fit model is shown by the blue solid line. The efficiencies based on the yields extracted from the fit are compared to the numbers obtained from the counted yields.

The trigger efficiency as a function of the probe muon's pseudorapidity is shown in Fig. 11.26. The integrated efficiency for at least one of the two muons to fire the Hlt1SingleMuonHighPT trigger is 0.972 ± 0.018 for pA and 0.967 ± 0.018 for Ap. The integration uses again the pseudorapidity spectrum of the muon from simulation at $\sqrt{s} = 5$ TeV.
The quoted systematic uncertainty is based on the binomial uncertainty on the efficiencies in each η bin and the uncertainty on the η spectrum determined in the same way as described in the tracking efficiency section. The two uncertainty components are propagated to the final value in the integration over the considered η range.
As a cross check the selection efficiency has been estimated with a wider ($m_{\mu^+\mu^-} \in [60, 120]\,\mathrm{GeV}/c^2$) and tighter ($m_{\mu^+\mu^-} \in [80, 100]\,\mathrm{GeV}/c^2$) invariant dimuon mass window, but the integrated efficiency values do not show any significant deviation.

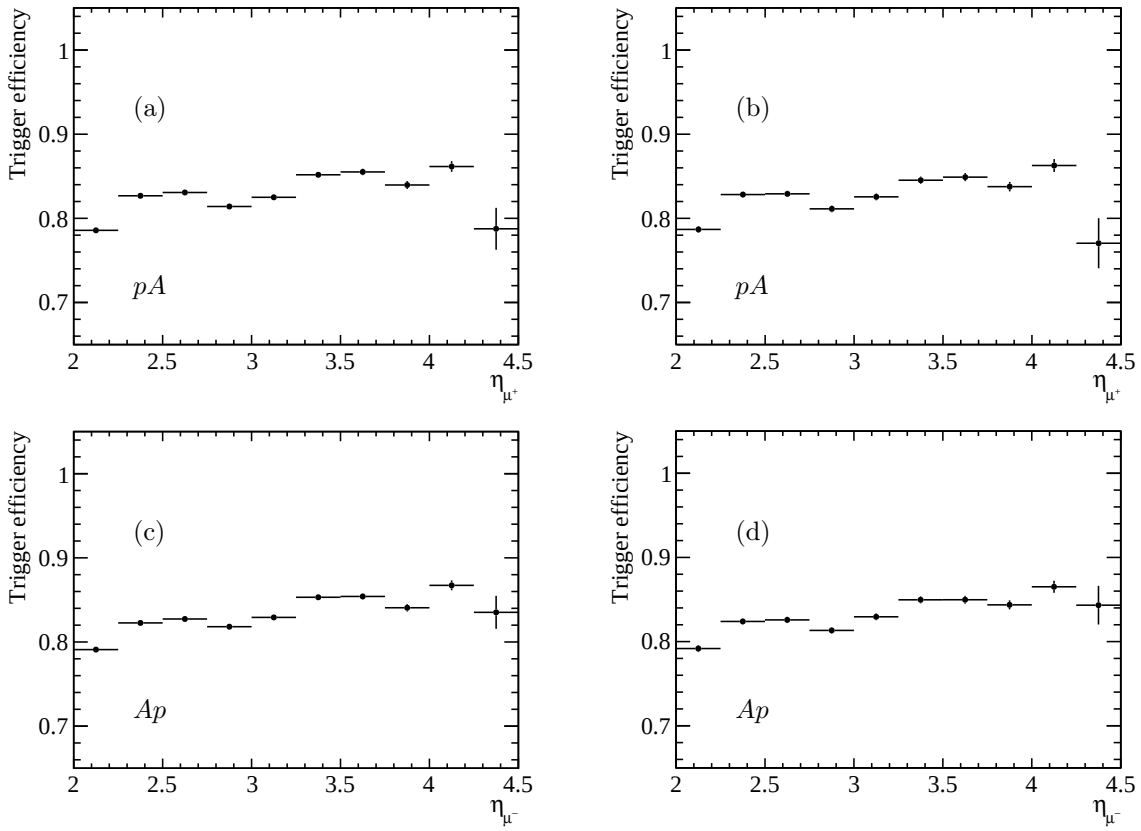


Figure 11.26: The estimated trigger efficiency as a function of the pseudorapidity of the track for (a) positive and (c) negative charged muons from $Z \to \mu^+\mu^-$ in pA data. Figures (b) and (d) show the same for positive and negative charged muons in Ap data.

11.4.5 Muon-Identification Efficiency $\varepsilon^{\mathbf{MUON|TRIG}}$

The muon-identification efficiency is estimated from a sample of $Z \to \mu^+\mu^-$ candidates in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ with the requirements of p_T larger than $20\,\mathrm{GeV}/c$ and η between 2.0 and 4.5, but no muon-identification requirements. The muon identification is required in the tag

sample for the tag muon and in the probe sample for both muons. Additional cuts are implied on the samples to reduce the background from random combinations of tracks as well as from combinations of a track and a corresponding ghost track in the tag sample: the absolute difference in the azimuthal angle $|\Delta\phi|$ between the two muons must be larger than 2.7 and the summed p_{T} of tracks clustered by the anti-k_t algorithm in a cone of $R = 0.5$ around the tag muon is required to be smaller than $2\,\mathrm{GeV}/c$. The muon-identification efficiency as a function of the η of the probe muon is shown in Fig. 11.27. The integrated and combined muon-identification efficiencies for both muons are 0.987 ± 0.023 for pA and 0.990 ± 0.021 for Ap.
The quoted systematic uncertainty is based on the binomial uncertainty on the efficiencies in each η bin and the uncertainty on the η spectrum determined in the same way as described in the tracking efficiency section. The two uncertainty components are propagated to the final value in the integration over the considered η range.
The muon-identification efficiency has been determined as a cross check using a tighter requirement of $|\Delta\phi| > 2.8$ showing no significant change of the integrated efficiency.
As a further cross check the selection efficiency has been estimated with a wider ($m_{\mu^+\mu^-} \in [60, 120]\,\mathrm{GeV}/c^2$) and tighter ($m_{\mu^+\mu^-} \in [80, 100]\,\mathrm{GeV}/c^2$) invariant dimuon mass window, but the integrated efficiency values do not show any significant deviation.

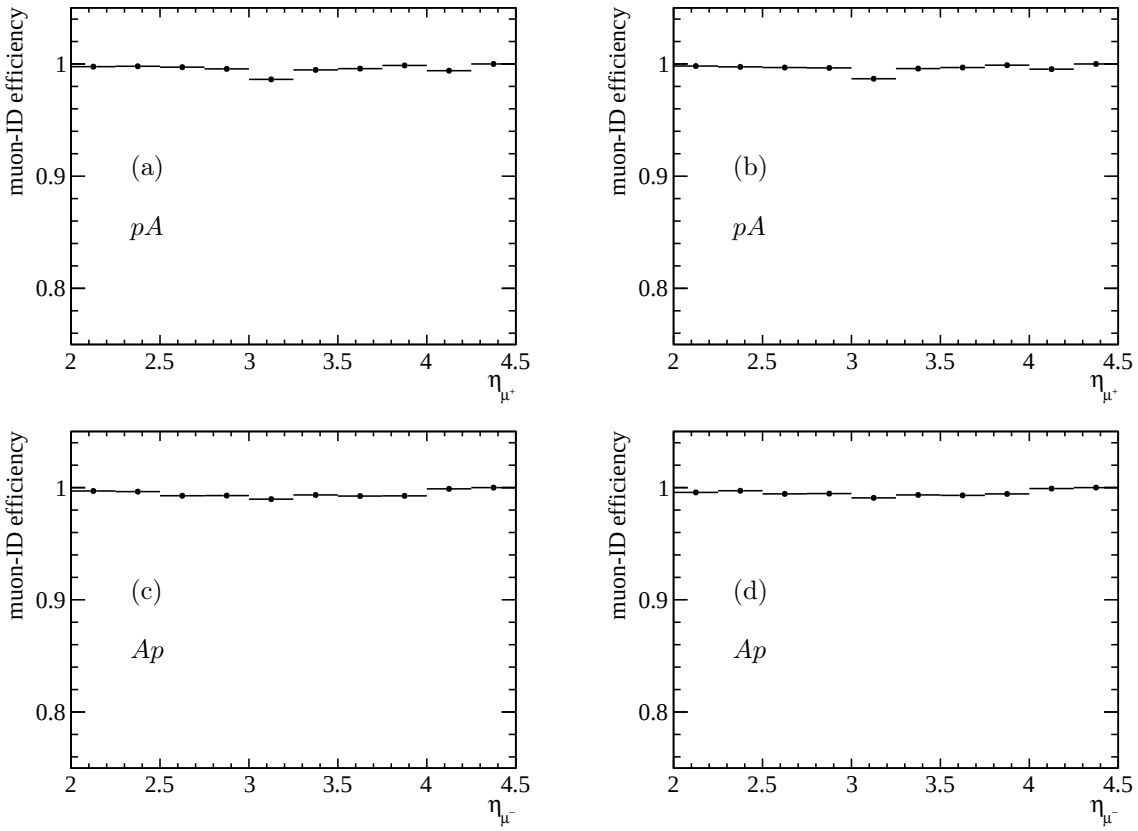


Figure 11.27: The estimated muon-id efficiency as a function of the pseudorapidity of the track for (a) positive and (c) negative charged muons from $Z \to \mu^+\mu^-$ in pA data. Figures (b) and (d) show the same for positive and negative charged muons in Ap data.

11.5 Systematic Uncertainties

There are systematic uncertainties assigned to the estimation of the purity, the GEC efficiency, the tracking, selection, trigger and muon-identification efficiency as well as the multiplicity reweighting.
The estimation of the systematic uncertainties as well as other checks that have been performed are described in the corresponding sections. Table 11.5 summarises the systematic uncertainties on the final cross section values. They are added in quadrature to give the total systematic uncertainty except the uncertainties of the reconstruction, selection, trigger and muon-identification efficiency. They are assumed to be fully correlated as they are estimated from the same data sample and using similar methods.

Table 11.5: Summary of the relative systematic uncertainties in the cross section calculation for $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$.

Quantity [a]	$\boldsymbol{pA}$	$\boldsymbol{Ap}$
Purity ρ	0.5 %	0.5 %
$\varepsilon^{\mathrm{GEC}}$	0.0 %	1.9 %
$\varepsilon^{\mathrm{REC}}$	2.0 %	2.1 %
$\varepsilon^{\mathrm{SEL\|REC}}$	2.2 %	2.6 %
$\varepsilon^{\mathrm{TRIG\|SEL}}$	1.9 %	1.9 %
$\varepsilon^{\mathrm{MUON\|TRIG}}$	2.3 %	2.1 %
Reweighting	1.5 %	2.0 %
Luminosity	1.9 %	2.1 %
Total	8.8 %	9.4 %

[a] As the reconstruction, selection, trigger and muon-id efficiencies have been estimated from the same data sample and with similar methods, a full correlation among these uncertainties is assumed.

11.6 Results

11.6.1 Cross Section Results

The quantities to estimate the cross section of $Z \to \mu^+\mu^-$ in pA and Ap collisions defined by Eq. (11.1) are summarised in Tab. 11.6.

The measured cross sections are

$$\begin{aligned} \sigma_{Z\to\mu^+\mu^-,pA} &= 13.5^{+5.4}_{-4.0}(\text{stat.}) \pm 1.1(\text{syst.}) \pm 0.3(\text{lumi.})\ \text{nb} \\ \sigma_{Z\to\mu^+\mu^-,Ap} &= 10.7^{+8.4}_{-5.1}(\text{stat.}) \pm 1.0(\text{syst.}) \pm 0.2(\text{lumi.})\ \text{nb} \end{aligned}$$

where the first uncertainty is the statistical uncertainty defined as 68 % confidence interval with symmetric coverage, the second one the systematic and the third one the uncertainty on the luminosity.

Table 11.6: Summary of the inputs to the cross section calculation for $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$.

Quantity	***pA***	***Ap***
N_{cand}	$11.0\ ^{+4.4}_{-3.3}$	$4.0\ ^{+3.1}_{-1.9}$
Purity ρ	0.997±0.005	0.996±0.005
Luminosity $\mathcal{L}$ [nb^{-1}]	1.099±0.021	0.521±0.011
ε^{GEC}	1.000±0.000	0.978±0.019
ε^{REC}	0.903±0.018	0.892±0.019
$\varepsilon^{\text{SEL\|REC}}$	0.855±0.019	0.853±0.022
$\varepsilon^{\text{TRIG\|SEL}}$	0.972±0.018	0.967±0.018
$\varepsilon^{\text{MUON\|TRIG}}$	0.987±0.023	0.990±0.021

The systematic uncertainties correspond to the propagated uncertainties on the purity and the efficiencies. A correlation of 1.0 is assumed among the tracking, trigger, selection and muon-identification efficiency.

The cross section for pA is – as shown by Fig. 11.28 – in very good agreement with the predictions discussed in the previous chapter while the observed value for Ap is by a factor of 4.1 larger than the predictions including nuclear effects, but considering the uncertainty of the measurement, this only corresponds to a deviation from the prediction by 1.5σ.

Statistical Significance of the Result

The statistical significance of the observed yields is calculated based on the number of candidates and the estimated purity. The p-value is given by

$$\begin{aligned} p &= \sum_{n=N_{\text{cand}}}^{\infty} \frac{\mu^n}{n!} e^{-\mu} \qquad (11.8) \\ \text{with } \mu = N_{\text{bkg,exp}} &= \sigma_{Z\to\mu^+\mu^-,\text{exp}} \cdot (1-\rho)/\rho \cdot \\ &\quad \mathcal{L} \cdot \varepsilon^{\text{GEC}} \cdot \varepsilon^{\text{REC}} \cdot \varepsilon^{\text{SEL|REC}} \cdot \varepsilon^{\text{TRIG|SEL}} \cdot \varepsilon^{\text{MUON|TRIG}}. \end{aligned}$$

The obtained p-value for pA is 1.9×10^{-25} and for Ap 6.9×10^{-12} corresponding to 10.4σ and 6.8σ, respectively when using the cross section prediction from FEWZ using EPS09. The numbers using the FEWZ predictions without taking nuclear modifications into account are 5.5×10^{-25} (10.3σ) in pA and 9.2×10^{-12} (6.7σ) in Ap.
In contrast to nucleus-nucleus collisions where the production of massive gauge bosons has already been observed in lead-lead collisions [4,5,11,110,113], there has not been so far any observation of the production of such bosons in proton-nucleus collisions. Therefore the measurement reported here is the first observation of the production of massive gauge bosons in proton-nuclei collisions.

11.6.2 Results for R_{FB}

As explained in Sec. 10.3.2, the ratio of the cross sections in the forward and backward direction, R_{FB}, is particularly sensitive to nuclear effects in the $Z \to \mu^+\mu^-$ production. The overlap

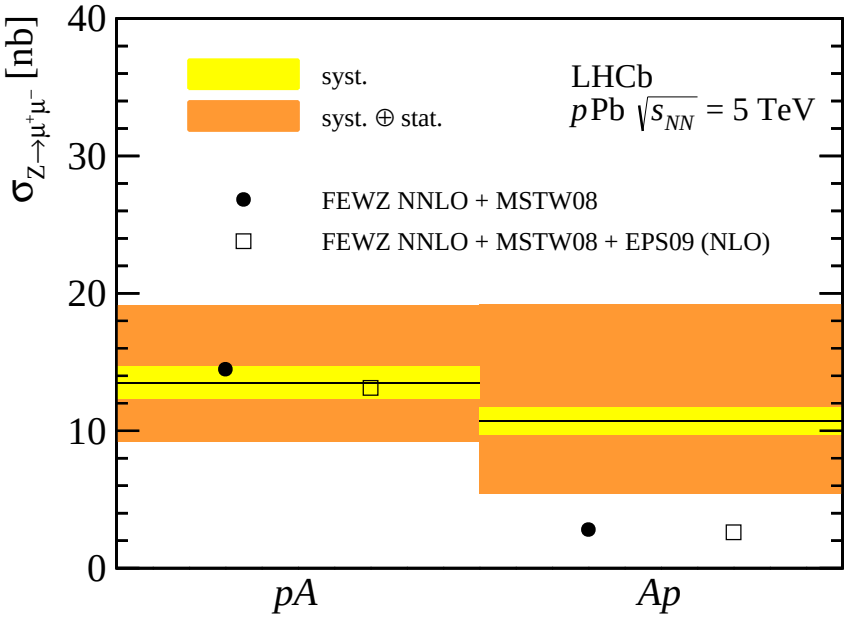


Figure 11.28: Summary plot of the results on $\sigma_{Z\to\mu^+\mu^-,pA/Ap}$ from LHCb with about 1.1 nb^{-1} of pA data and 0.5 nb^{-1} of Ap data. The black line shows the measured values with the systematic uncertainty displayed by the yellow and the statistical and systematic uncertainties by the orange band. The black points show the predictions based on NNLO calculations using FEWZ with MSTW08 with no nuclear modifications and the hollow one the predictions based on NNLO calculations using FEWZ with MSTW08 and EPS09 at NLO. The scale and PDF uncertainties on the predictions are – compared to the uncertainties of the measurement – very small and thus not visible.

between the pA and Ap samples expressed in terms of the absolute rapidity of the Z boson in the centre-of-mass frame, $|y|$, ranges from 2.5 to 4.0.
R_{FB} is calculated as

$$\begin{aligned} R_{\mathrm{FB}}(2.5<|y|<4.0) &= \frac{\sigma_{Z\to\mu^+\mu^-,pA}(2.5<|y|<4.0)}{\sigma_{Z\to\mu^+\mu^-,Ap}(2.5<|y|<4.0)}\cdot\beta^{-1} \\ &= \frac{N_{\mathrm{cand},pA}}{N_{\mathrm{cand},Ap}}\Big|_{2.5<|y|<4.0}\cdot\frac{\mathcal{L}_{Ap}}{\mathcal{L}_{pA}}\cdot \\ &\quad \frac{\varepsilon^{\mathrm{GEC}}_{Ap}\cdot\varepsilon^{\mathrm{REC}}_{Ap}\cdot\varepsilon^{\mathrm{SEL|REC}}_{Ap}\cdot\varepsilon^{\mathrm{TRIG|SEL}}_{Ap}\cdot\varepsilon^{\mathrm{MUON|TRIG}}_{Ap}}{\varepsilon^{\mathrm{GEC}}_{pA}\cdot\varepsilon^{\mathrm{REC}}_{pA}\cdot\varepsilon^{\mathrm{SEL|REC}}_{pA}\cdot\varepsilon^{\mathrm{TRIG|SEL}}_{pA}\cdot\varepsilon^{\mathrm{MUON|TRIG}}_{pA}}\cdot\beta^{-1}, \end{aligned} \tag{11.9}$$

where β is the acceptance correction as the cuts on η and p_{T} have a different impact in the two configurations. β is estimated from FEWZ NNLO calculations giving

$$\beta = 2.419^{+0.127}_{-0.000}(\mathrm{theo.}) \pm 0.008(\mathrm{num.})^{+0.009}_{-0.010}(\mathrm{PDF}),$$

where the first uncertainty comes from a variation of the factorisation and renormalisation scale by a factor two up- and downward. The second uncertainty is due to the numerical precision of the calculation and the third one comes from the uncertainty on the PDF set.
The efficiencies are not recalculated for Z candidates with $2.5<|y|<4.0$ as the total efficiency as a function of the muon pseudorapidity is close to flat.

In the case of the Ap configuration, all four candidates fall in this range while for pA only two candidates have a rapidity high enough, so the ratio $r = N_{\mathrm{cand},pA}/N_{\mathrm{cand},Ap}|_{2.5<|y|<4.0}$ is equal to 0.5. The statistical uncertainty on this ratio is defined as 68 % confidence interval $[r_{\mathrm{low}}, r_{\mathrm{high}}]$ with symmetric coverage and is calculated according to

$$\frac{\int_r^\infty \mathrm{d}r' \int\int_{r_{\mathrm{low}}=\lambda_1/\lambda_2} \mathrm{d}\lambda_1\mathrm{d}\lambda_2 p(r'|\lambda_1,\lambda_2)\cdot\mathcal{L}(\lambda_1,\lambda_2|N_{pA},N_{Ap})}{\int\int_{r_{\mathrm{low}}=\lambda_1/\lambda_2}\mathrm{d}\lambda_1\mathrm{d}\lambda_2\mathcal{L}(\lambda_1,\lambda_2|N_{pA},N_{Ap})} = \frac{1-\alpha}{2} \tag{11.10}$$

$$\frac{\int_0^r \mathrm{d}r' \int\int_{r_{\mathrm{high}}=\lambda_1/\lambda_2} \mathrm{d}\lambda_1\mathrm{d}\lambda_2 p(r'|\lambda_1,\lambda_2)\cdot\mathcal{L}(\lambda_1,\lambda_2|N_{pA},N_{Ap})}{\int\int_{r_{\mathrm{high}}=\lambda_1/\lambda_2}\mathrm{d}\lambda_1\mathrm{d}\lambda_2\mathcal{L}(\lambda_1,\lambda_2|N_{pA},N_{Ap})} = \frac{1-\alpha}{2} \tag{11.11}$$

with

$$\mathcal{L}(\lambda_1,\lambda_2|N_{pA},N_{Ap}) = \frac{\lambda_1^{N_{pA}}}{N_{pA}!}e^{-\lambda_1}\cdot\frac{\lambda_2^{N_{Ap}}}{N_{Ap}!}e^{-\lambda_2}$$

and α is 0.683. $p(r|\lambda_1,\lambda_2)$ is the distribution of the ratio of two Poisson variables with expectation values λ_1 and λ_2 and is sampled by using Toy MC.
The resulting confidence interval (*cf.* Fig. 11.29(a)) gives the following result:

$$r = \frac{N_{\mathrm{cand},pA}}{N_{\mathrm{cand},Ap}}\Big|_{2.5<|y|<4.0} = 0.50^{+0.55}_{-0.33}$$

The systematic uncertainty on R_{FB} is estimated by the usual error propagation of the uncertainties on the different efficiencies and the purity as well as the acceptance correction. As the purity and efficiencies for pA and Ap are estimated from the same samples and with the same technique, a correlation between the corresponding quantities for pA and Ap of 1.0 is assumed.

The uncertainty of the luminosities for pA and Ap is uncorrelated as the main component of the uncertainties is statistical and there is only a very small common uncertainty due to the method [9].

The final value of R_{FB} is

$$R_{\mathrm{FB}}(2.5 < |y| < 4.0) = 0.094^{+0.104}_{-0.062}(\mathrm{stat.})^{+0.002}_{-0.006}(\mathrm{syst.}) \pm 0.003(\mathrm{lumi.}).$$

The quoted statistical uncertainties correspond to the 68 % confidence interval with symmetric coverage. The 99.7 % (*i.e.* 3σ) confidence interval with symmetric coverage is $[0.002, 1.626]$ whereas the asymmetry of the interval around the central value is due to non-Gaussian statistical uncertainties of R_{FB}.

Statistical Significance of R_{FB}

The result on R_{FB} is compared to the null-hypothesis of $R_{\mathrm{FB}} = 1.0$ by constructing the distribution of R_{FB} assuming a true value of 1.0 with Toy Monte Carlo in the same manner as described above for the yield ratio. In addition, the combined systematic and luminosity uncertainty – although much smaller than the statistical one – are additionally taken into account by a convolution of the distribution and a Gaussian with a corresponding standard deviation. The resulting distribution is shown in Fig. 11.29(b).
The p-value estimated from this distribution for $R_{\mathrm{FB}} = 1.0$ is 1.2 %, which corresponds to 2.2σ. So in summary, the obtained value for R_{FB} is, as expected, lower than one, but does not show a significant deviation from one due to the large statistical uncertainty.
The downward fluctuation is driven by two effects: first the measured cross section for Ap is larger than expected and second the observed distribution in y for the pA data is shifted towards lower values with respect to the expected one (*cf.* Fig. 11.5).

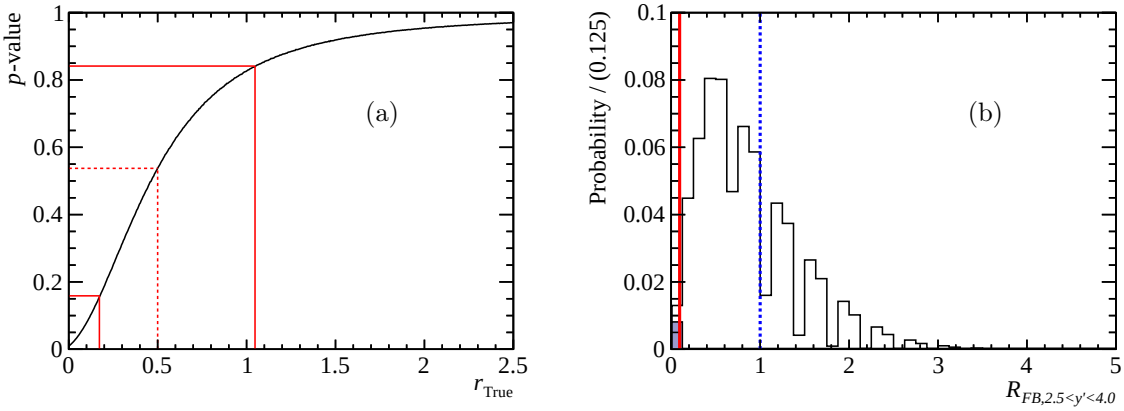


Figure 11.29: (a) The integral of the cumulative distribution function of r for the observed value of r as a function of the true value of r. The solid lines indicate the lower and upper boundary of the 68 % confidence interval; (b) The distribution for R_{FB} for a true value of 1.0 from Toy Monte Carlo. The solid red line shows the central value and the dashed blue one $R_{\mathrm{FB}} = 1.0$. The filled part of the histogram shows the region below the observed value of R_{FB} corresponding to a p-value of 0.012. The discrete behaviour of the plots originates from the fact that the ratio is built from integers.

Conclusion

Between spring 2010 and spring 2013 the LHC and its experiments performed their first period of data taking, called Run I. This data taking period has been a great success for the LHCb experiment. The LHCb detector showed an excellent performance and was able to cope with data taking conditions that led to an instantaneous luminosity twice as high as the design value. This thesis discussed the measurement of the radiation damage in the LHCb Silicon Tracker with two different methods, based on the measured leakage currents and on Charge Collection Efficiency scans using dedicated collision data. The measured evolution of the radiation damage in the Silicon Tracker shows good agreement with expectations and is an example of the good predictive power of the radiation simulations used in the design of LHCb.
In terms of physics analysis, the LHCb experiment has performed a wide range of measurements in different fields of high energy physics, not only in its main field of Flavour Physics. One of the highlights from Run I has been the first evidence for the very rare decay $B_s^0 \to \mu^+\mu^-$. The measured branching fraction of the decay, based on the full data set collected in Run I is

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.9^{+1.2}_{-1.2}) \times 10^{-9},$$

where the quoted uncertainty is the statistical and systematic uncertainties added in quadrature. This value is in good agreement with the time-averaged Standard Model (SM) prediction of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)^{\langle t \rangle} = (3.56 \pm 0.30) \times 10^{-9}$. Therefore, the result allows to rule out models of physics beyond the SM that predict a significant enhancement or reduction of the branching fraction with respect to the SM.
The measured branching fraction of the decay $B^0 \to \mu^+\mu^-$ is

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.7^{+2.5}_{-2.1}) \times 10^{-10},$$

where the quoted uncertainty is also the combined statistical and systematic uncertainties. The SM prediction for the branching fraction of this decay is $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.07 \pm 0.10)^{-10}$, which means that there is an enhancement in the measured value. However, within the collected event statistics, the tension between the measurement and the SM prediction is far from significant, and more data are needed to investigate it further.
This thesis described for the search of $B^0_{(s)} \to \mu^+\mu^-$ the techniques that were employed to estimate the expected distribution of signal decays in the invariant dimuon mass and the multivariate classifier based on topological and kinematical variables to separate the signal decays from background processes. The calibration of these distributions has been performed on data using dimuon resonances such as $J/\psi \to \mu^+\mu^-$ or $\Upsilon(1S) \to \mu^+\mu^-$ and charmless two-body decays of B^0 and B_s^0 mesons into pairs of kaons and pions ($K^\pm\pi^\mp, \pi^+\pi^-, K^+K^-$). The approach to use

data instead of simulation to calibrate the signal distribution allows to avoid possible biases due to imperfections in the simulation.
LHCb also performed measurements outside its main field of Flavour Physics, for example to study Quantum Chromodynamics and electroweak boson production or to search directly for exotic particles. One example is the first observation of Z production in proton-nucleus collisions. This measurement by LHCb using data collected during a short period of proton-lead runs of the LHC is described in this thesis. The $Z \to \mu^+\mu^-$ production cross section has been measured in the fiducial region defined by the cuts $p_{\mathrm{T}}(\mu) > 20\,\mathrm{GeV}/c$, $2.0 < \eta(\mu) < 4.5$ and $60 < m_{\mu^+\mu^-} < 120\,\mathrm{GeV}/c^2$, where the variables are defined in the laboratory frame. The obtained values are

$$\sigma_{Z\to\mu^+\mu^-,pA} = 13.5^{+5.1}_{-4.0}(\mathrm{stat.}) \pm 1.1(\mathrm{syst.}) \pm 0.3(\mathrm{lumi.})\ \mathrm{nb},$$

measured in the direction of the proton beam, and

$$\sigma_{Z\to\mu^+\mu^-,Ap} = 10.7^{+8.4}_{-5.1}(\mathrm{stat.}) \pm 1.0(\mathrm{syst.}) \pm 0.2(\mathrm{lumi.})\ \mathrm{nb},$$

measured in the direction of the lead beam. The results are limited by the available statistics, but provide a starting point for extended studies of electroweak boson production in proton-nucleus collisions, which could lead to a better understanding of the dynamics of partons inside nuclei.
For the upcoming Run II of LHC, an increase of the LHCb data sample by 5 fb^{-1} is expected. Due to the higher centre-of-mass energy of 13 to 14 TeV, the expected production cross section for beauty quarks or electroweak gauge bosons will significantly increase, approximately linearly as a function of the centre-of-mass energy. Thus these data samples will allow to obtain "more and more precise [measurements]", which will for example shed more light on the question whether the present excess in the $B^0 \to \mu^+\mu^-$ decay channel is due to physics beyond the SM or is just a statistical fluctuation.

Appendices

A

Additional Material on Radiation Damage Studies

The following pages show a set of figures giving further details on the CCE scan based radiation damage studies.
Figures A.1 and A.2 show the fitted signal height distributions for the different values of V_{bias}, with no sampling time shift applied (*i.e.* $\delta t = 0$ ns), in the TT and IT, respectively. The displayed distributions are from read-out sector 2642 in the CCE scan from September, 28 2013 for the TT and read-out sector 7300 in the same CCE scan for the IT.
Figures A.3 and A.4 show the pulse shape (*i.e.* signal height as a function of the signal sampling time shift δt) for the different voltage steps in the TT and IT, respectively. The shapes for the lowest value of V_{depl} have large uncertainties and are thus excluded from the determination of V_{depl}.
Figrues A.5, A.6 and A.7 show further examples of the measured depletion voltage as a function of time compared with the predictions from the full Hamburg model for several TT read-out sectors.
Figrue A.8 shows the change in the measured depletion voltage, ΔV_{depl}, in the innermost part of the TT as, with respect to the first CCE scan.

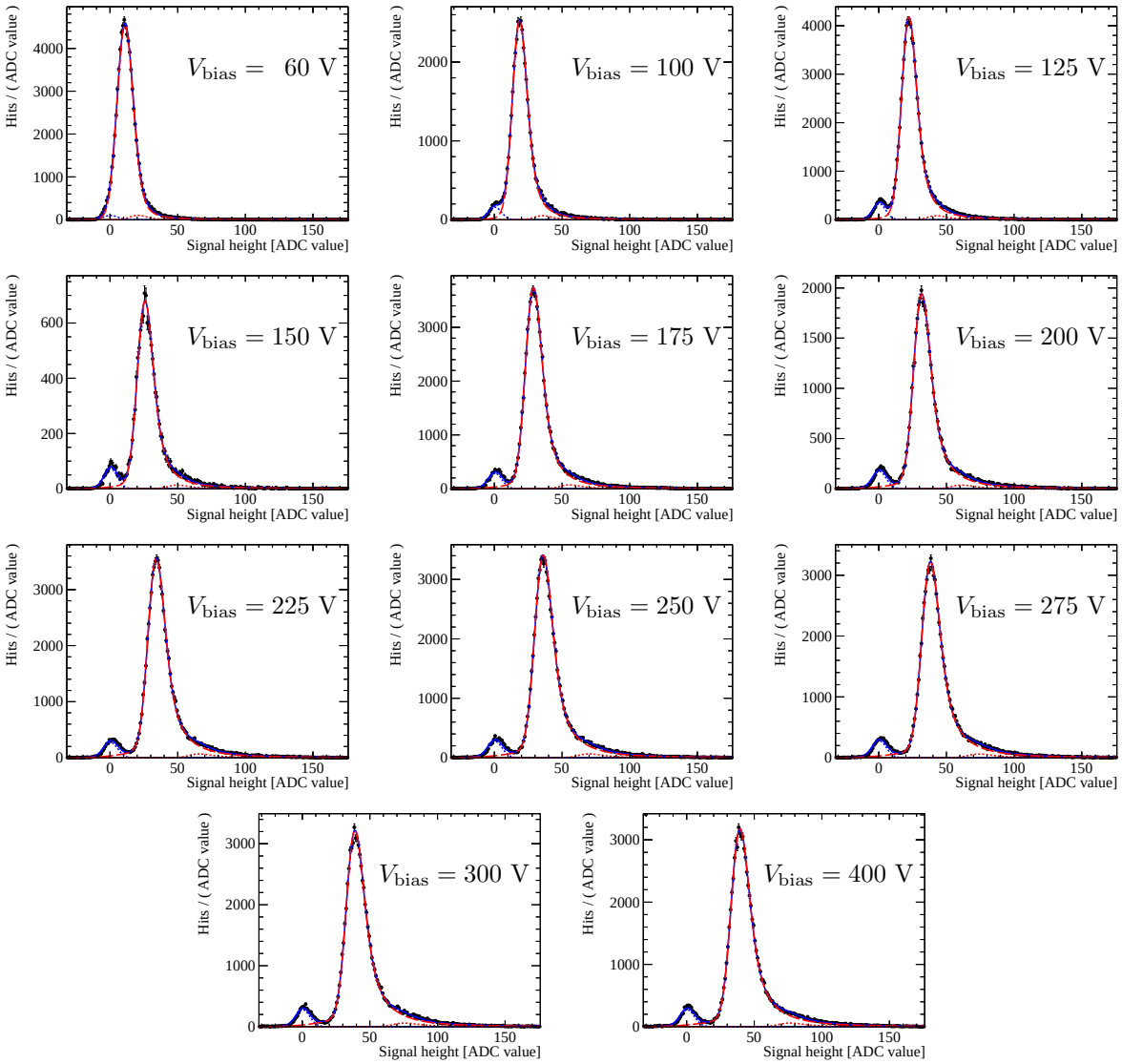


Figure A.1: ADC value distributions in the CCE scan from September, 28 2013 for the TT read-out sector 2642 with no sampling time shift (*i.e.* $\delta t = 0$ ns). The data distribution is fit by a double Gaussian for background with parameters fixed from a dedicated background sample. The distribution for the signal component is a convolution of a sum of two Landau distributions with the same double Gaussian. The second Landau distribution describing signals from converted photons has a most probable value and a width twice as high as the first and dominating Landau distribution.

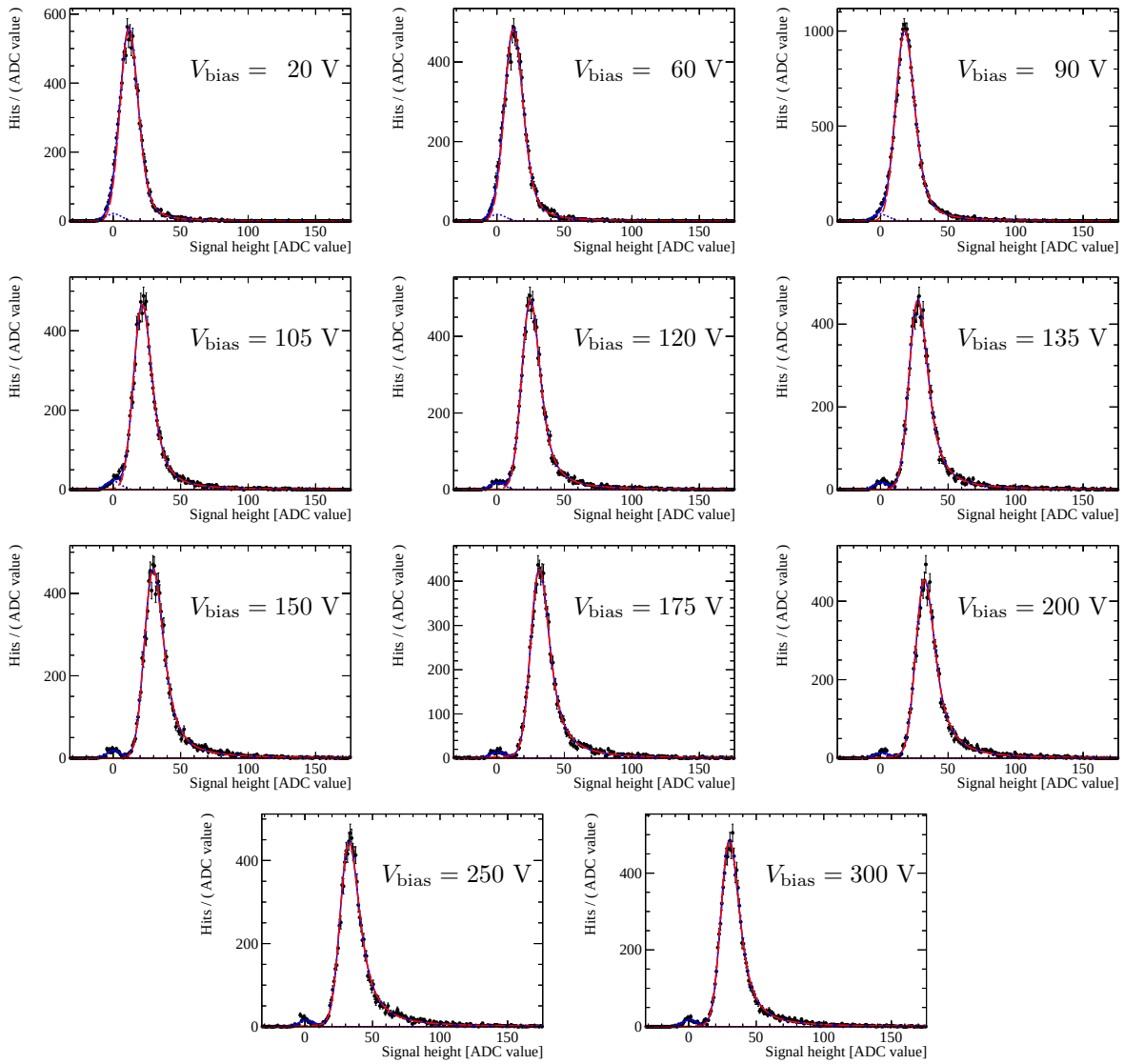


Figure A.2: ADC value distributions in the CCE scan from September, 28 2013 for the IT read-out sector 7300 with no sampling time shift (*i.e.* $\delta t = 0$ ns). The data distribution is fit by a double Gaussian for background with parameters fixed from a dedicated background sample. The distribution for the signal component is a convolution of a Landau distribution with the same double Gaussian.

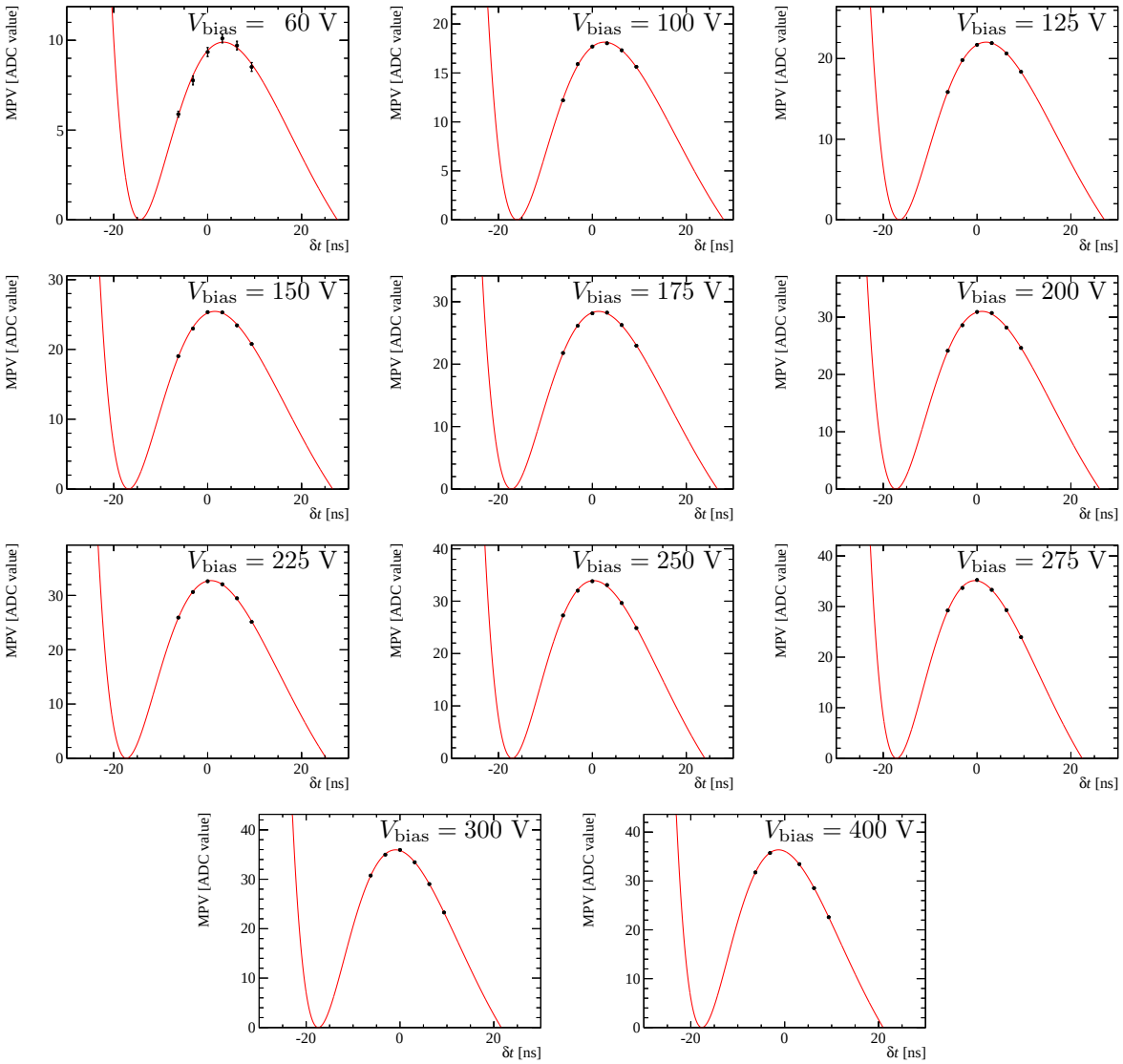


Figure A.3: Pulse shapes in the CCE scan from September, 28 2013 for the TT read-out sector 2642. The data are fit with the pulse shape form corresponding to the used CR-RC shaper in the ST read-out electronics. The uncertainties of the data points at the lowest value of V_{depl} lead to a large uncertainty on the pulse shape itself. Thus this V_{depl} value is discarded.

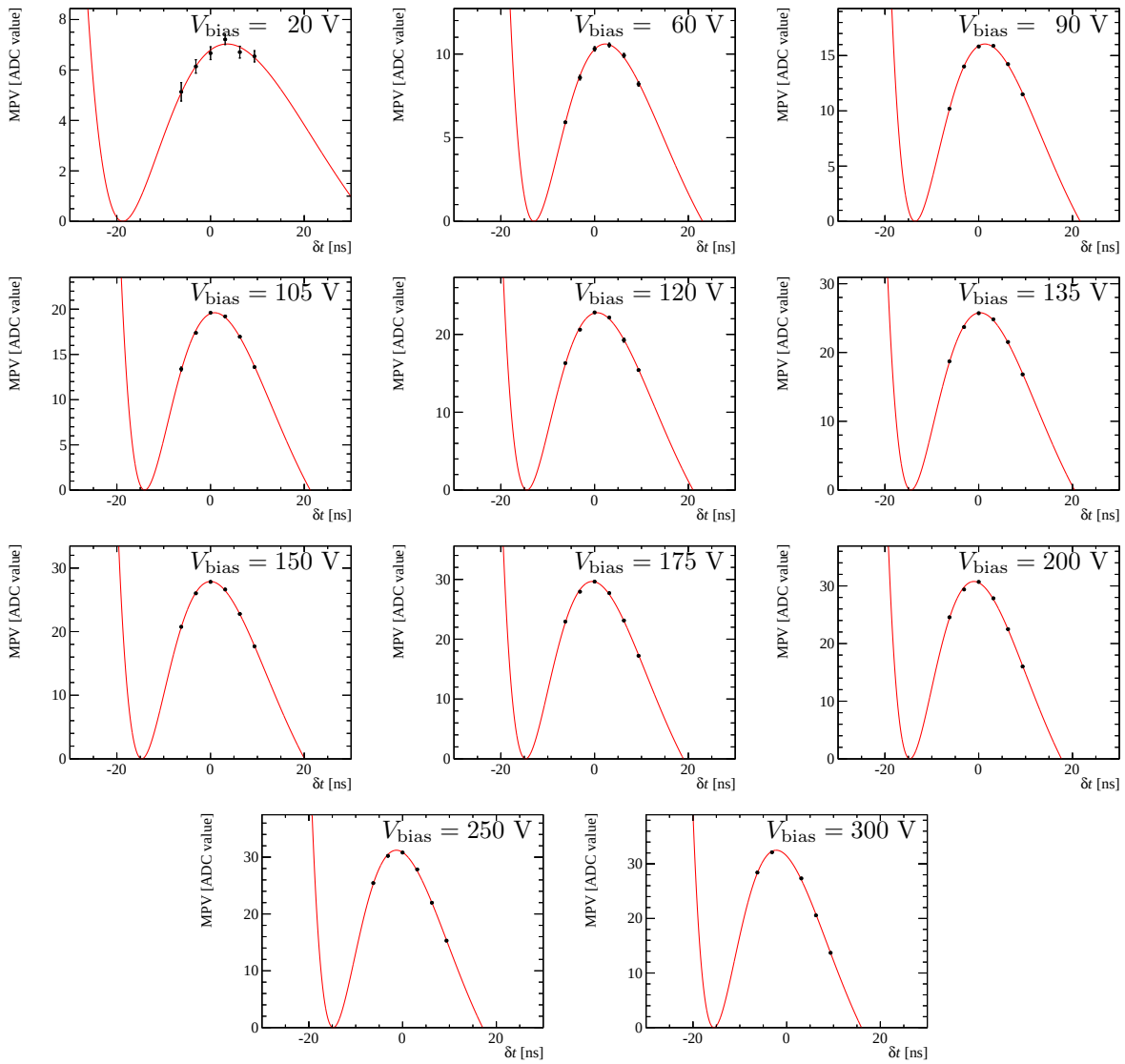


Figure A.4: Pulse shapes in the CCE scan from September, 28 2013 for the IT read-out sector 7300. The data are fit with the pulse shape form corresponding to the used CR-RC shaper in the ST read-out electronics. The uncertainties of the data points at the lowest value of V_{depl} lead to a large uncertainty on the pulse shape itself. Thus this V_{depl} value is discarded.

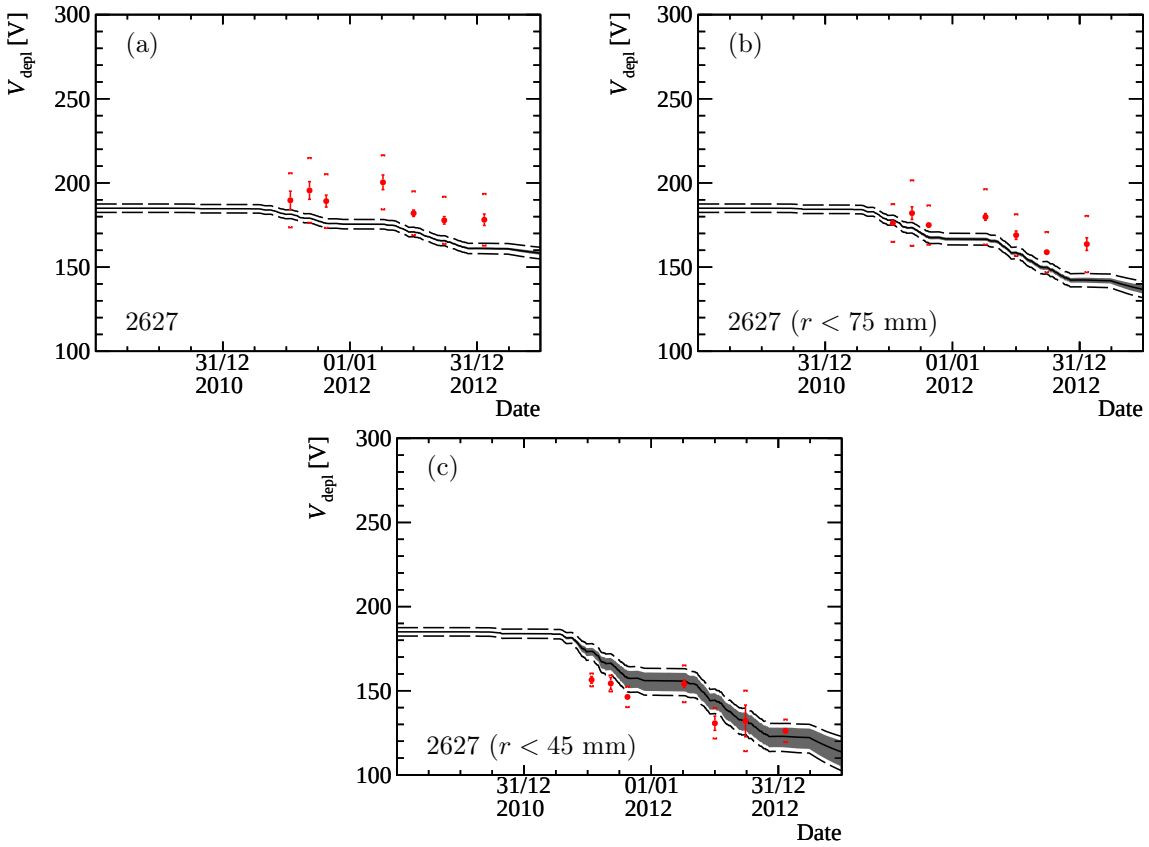


Figure A.5: V_{depl} as a function of time for the TT read-out sector 2627 ((a) full sensor, (b) region closer than 75 mm to the beam axis, (c) region closer than 45 mm to the beam axis). The red data points show the measurements from the CCE scan. The inner error bars correspond to the statistical, the outer to the systematic uncertainty. The black line shows the prediction for the Hamburg model with the grey band showing its uncertainty based on the uncertainties of the model parameter, of the FLUKA simulation and of the temperature measurement. The dashed lines show the ±5 V uncertainty on the initial V_{depl} measurement after production.

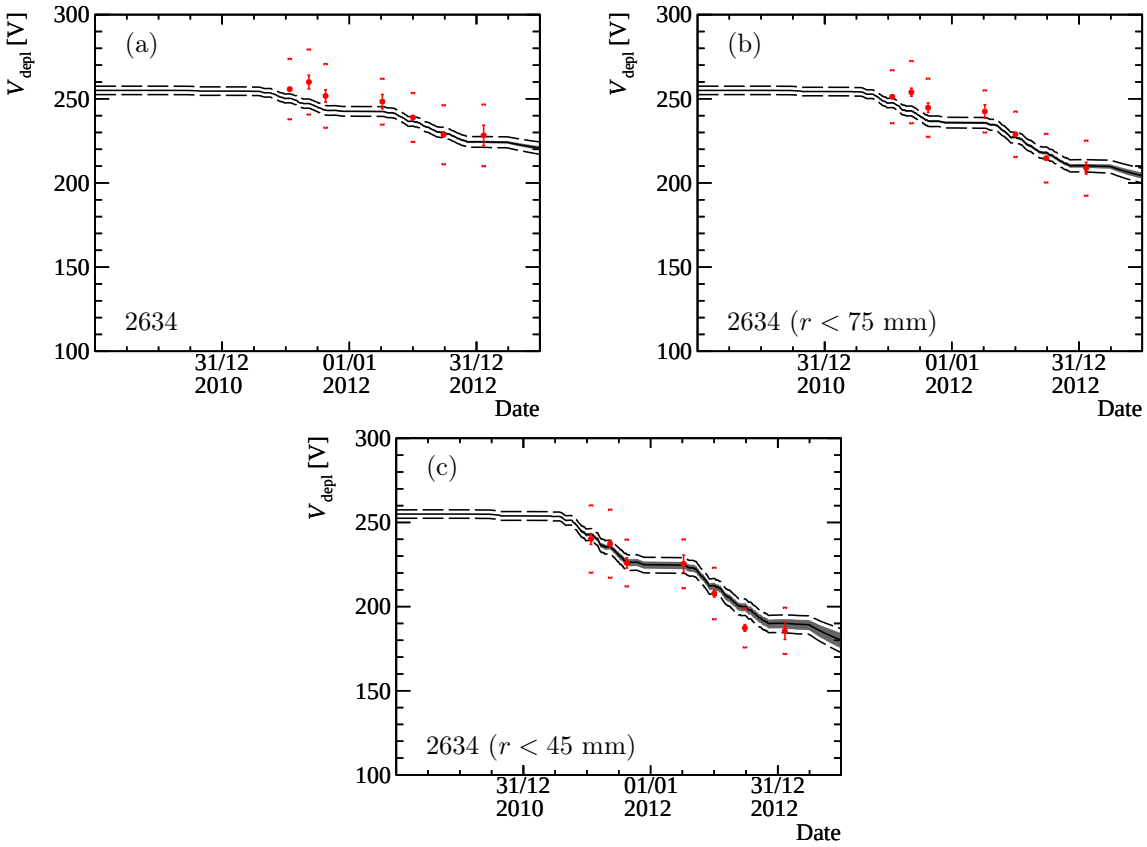


Figure A.6: V_{depl} as a function of time for the TT read-out sector 2634 ((a) full sensor, (b) region closer than 75 mm to the beam axis, (c) region closer than 45 mm to the beam axis). The red data points show the measurements from the CCE scan. The inner error bars correspond to the statistical, the outer to the systematic uncertainty. The black line shows the prediction for the Hamburg model with the grey band showing its uncertainty based on the uncertainties of the model parameter, of the FLUKA simulation and of the temperature measurement. The dashed lines show the ±5 V uncertainty on the initial V_{depl} measurement after production.

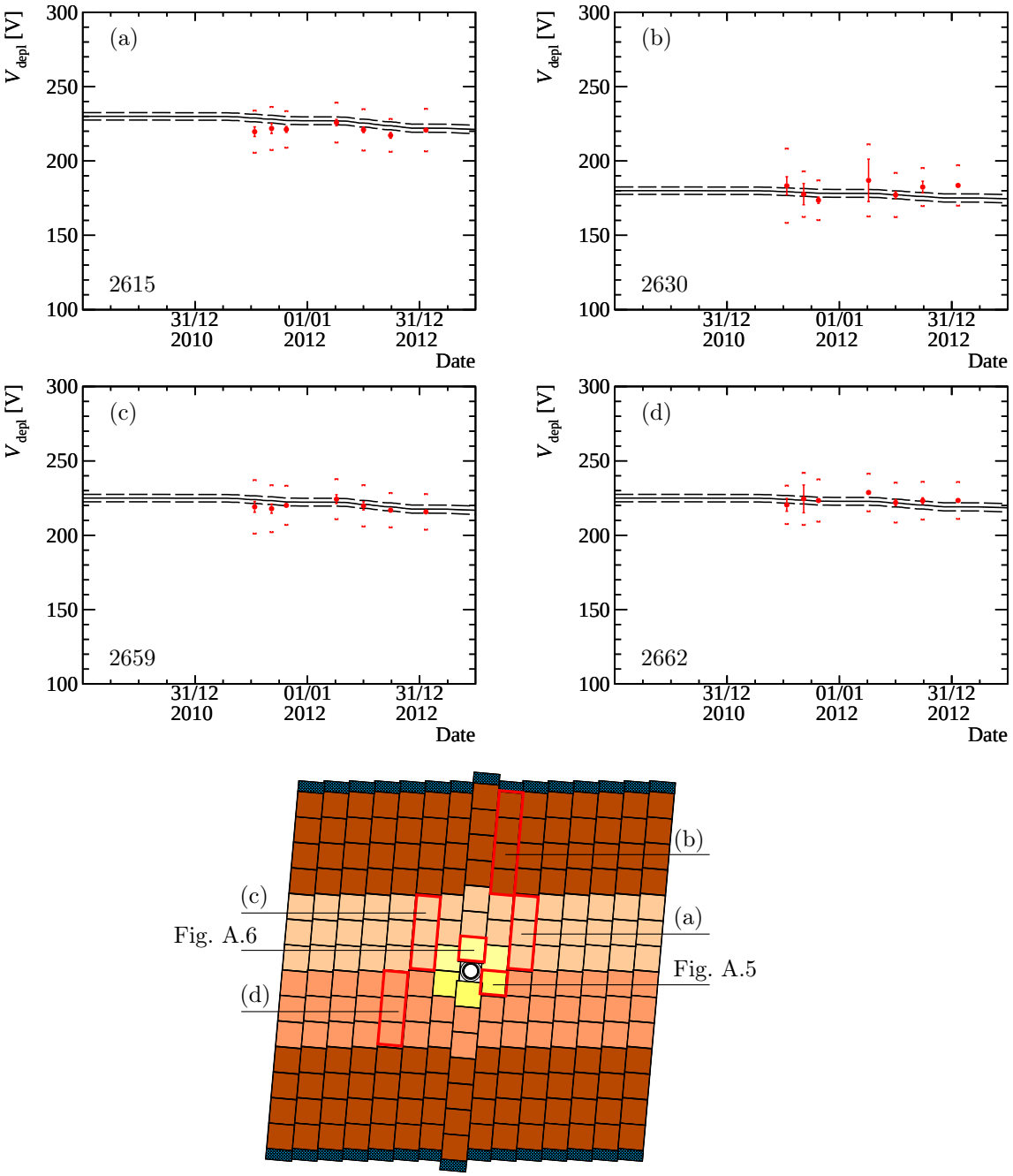
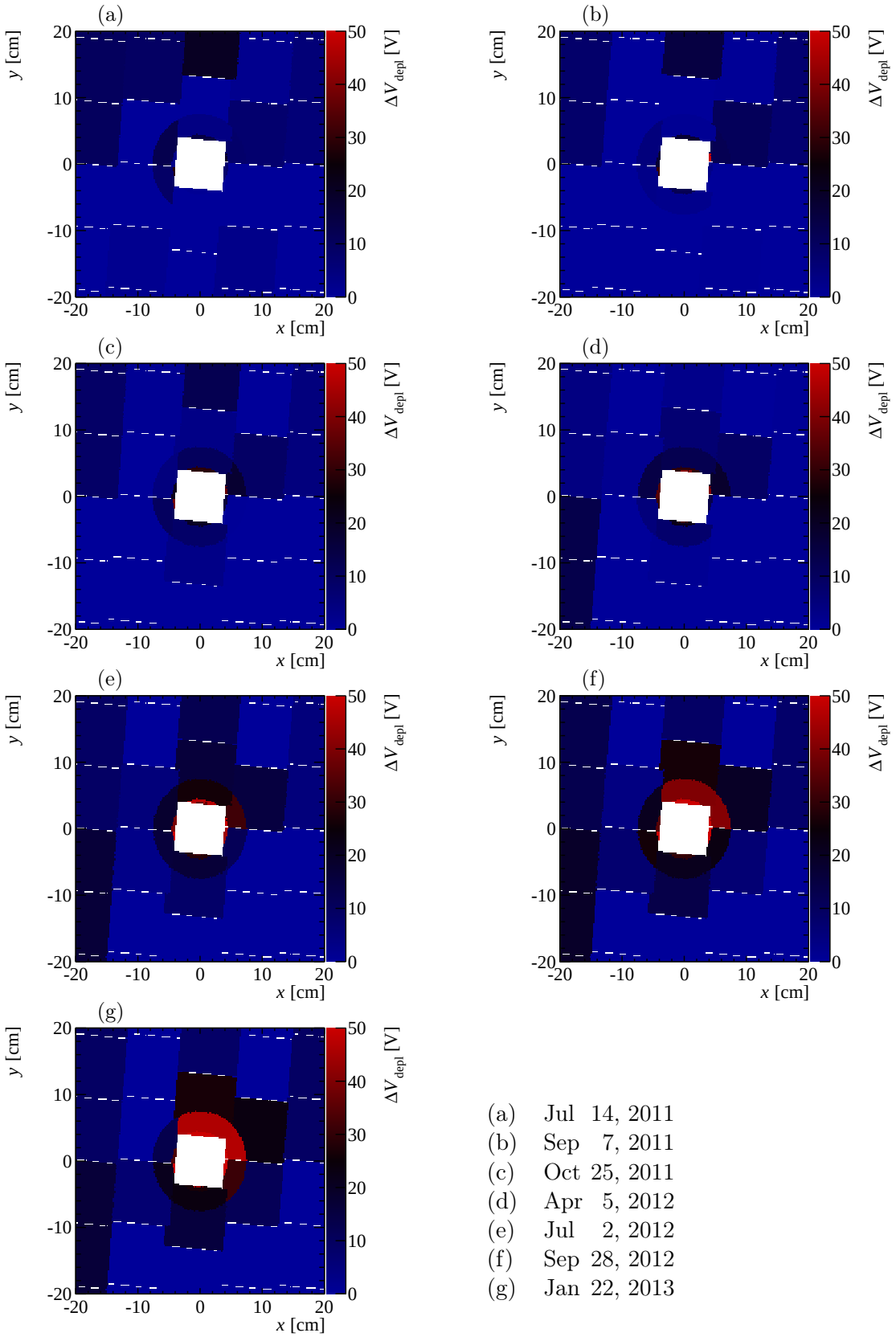


Figure A.7: V_{depl} as a function of time for the TT read-out sectors (a) 2615, (b) 2630, (c) 2659 and (d) 2662. The red data points show the measurements from the CCE scan. The inner error bars correspond to the statistical, the outer to the systematic uncertainty. The black line shows the prediction for the Hamburg model with the grey area showing its uncertainty based on the uncertainties of the model parameter, of the FLUKA simulation and of the temperature measurement. The dashed lines show the ± 5 V uncertainty on the initial V_{depl} measurement after production. The schematic at the bottom shows the position of the considered sectors, also for Figs A.5 and A.6.



(a) Jul 14, 2011
(b) Sep 7, 2011
(c) Oct 25, 2011
(d) Apr 5, 2012
(e) Jul 2, 2012
(f) Sep 28, 2012
(g) Jan 22, 2013

Figure A.8: Measured change in the depletion voltage, ΔV_{depl}, in the innermost part of TTaU with respect to the CCE scan in March 2011.

B

Tail Parameter Distributions in $B^0_{(s)} \to \mu^+\mu^-$

The following pages show details on the tail parameters of single and double Crystal Ball functions to describe the invariant dimuon mass distribution in $B^0_{(s)} \to \mu^+\mu^-$.
Figure B.1 shows the distribution of the tail parameters α and n for a single Crystal Ball function describing the invariant dimuon mass distribution of $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in 2011 data and 2012 data. Figures B.2, B.3, B.4 and B.5 show the tail parameters $\alpha_{l,r}$ and $n_{l,r}$ as well as f for a double Crystal Ball function describing the invariant dimuon mass distribution of $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in 2011 data and 2012 data.
These distributions reflect the variation of the parameters due to the uncertainty on the width, σ_{smear} of the Gaussian convolved with the distribution of the invariant dimuon mass calculated from the true four-momenta of the muons (single Crystal Ball function) or on the scale parameter, β to correct for the measured mass resolution in data (double Crystal Ball function).

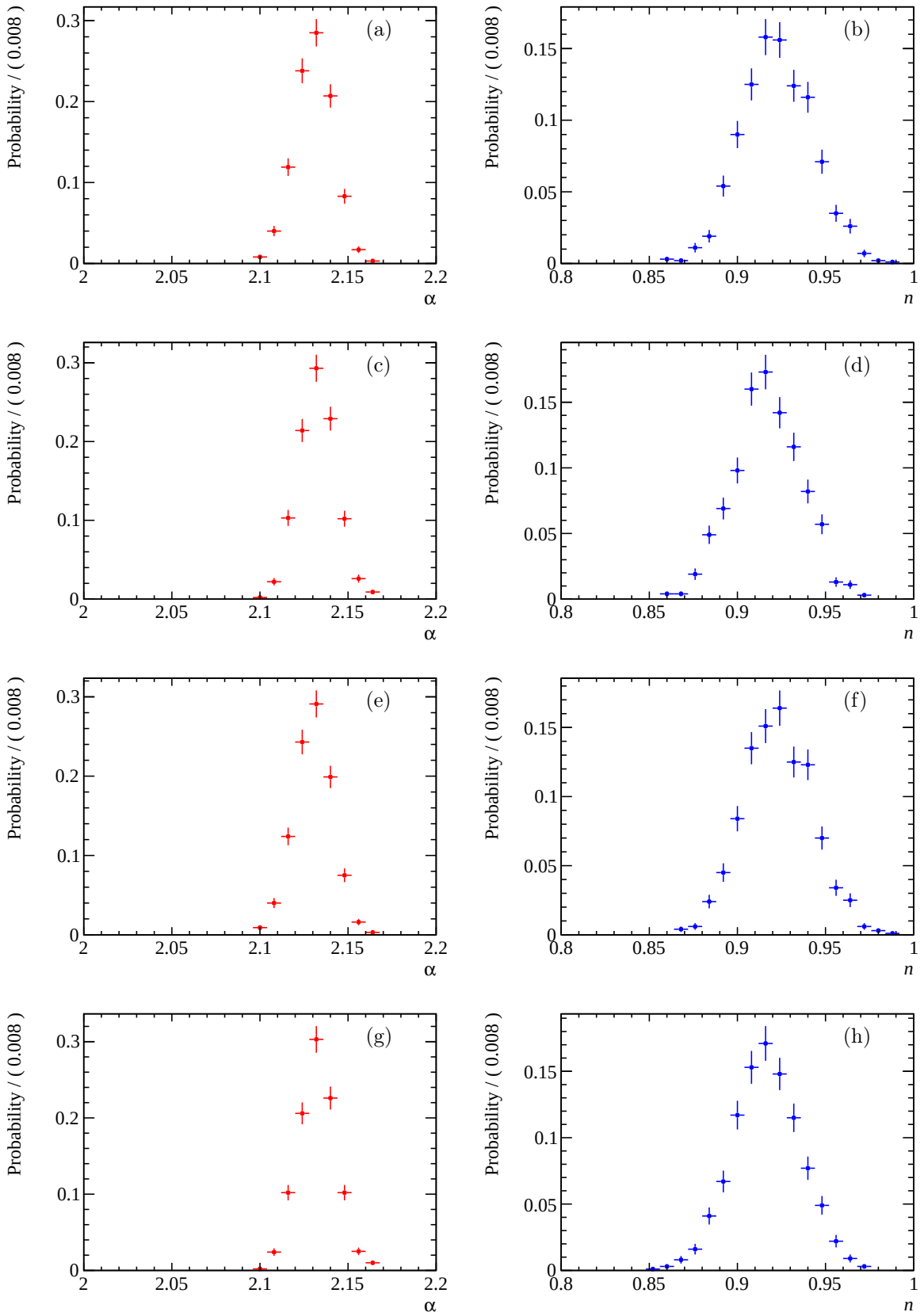


Figure B.1: The distribution of (a) α and (b) n for $B^0_s \to \mu^+\mu^-$ and (c) α and (d) n for $B^0 \to \mu^+\mu^-$ in 2011 data and (e) α and (f) n for $B^0_s \to \mu^+\mu^-$ and (g) α and (h) n for $B^0 \to \mu^+\mu^-$ in 2011 data. The standard deviation of the distributions is used as systematic uncertainty on the values of α and n.

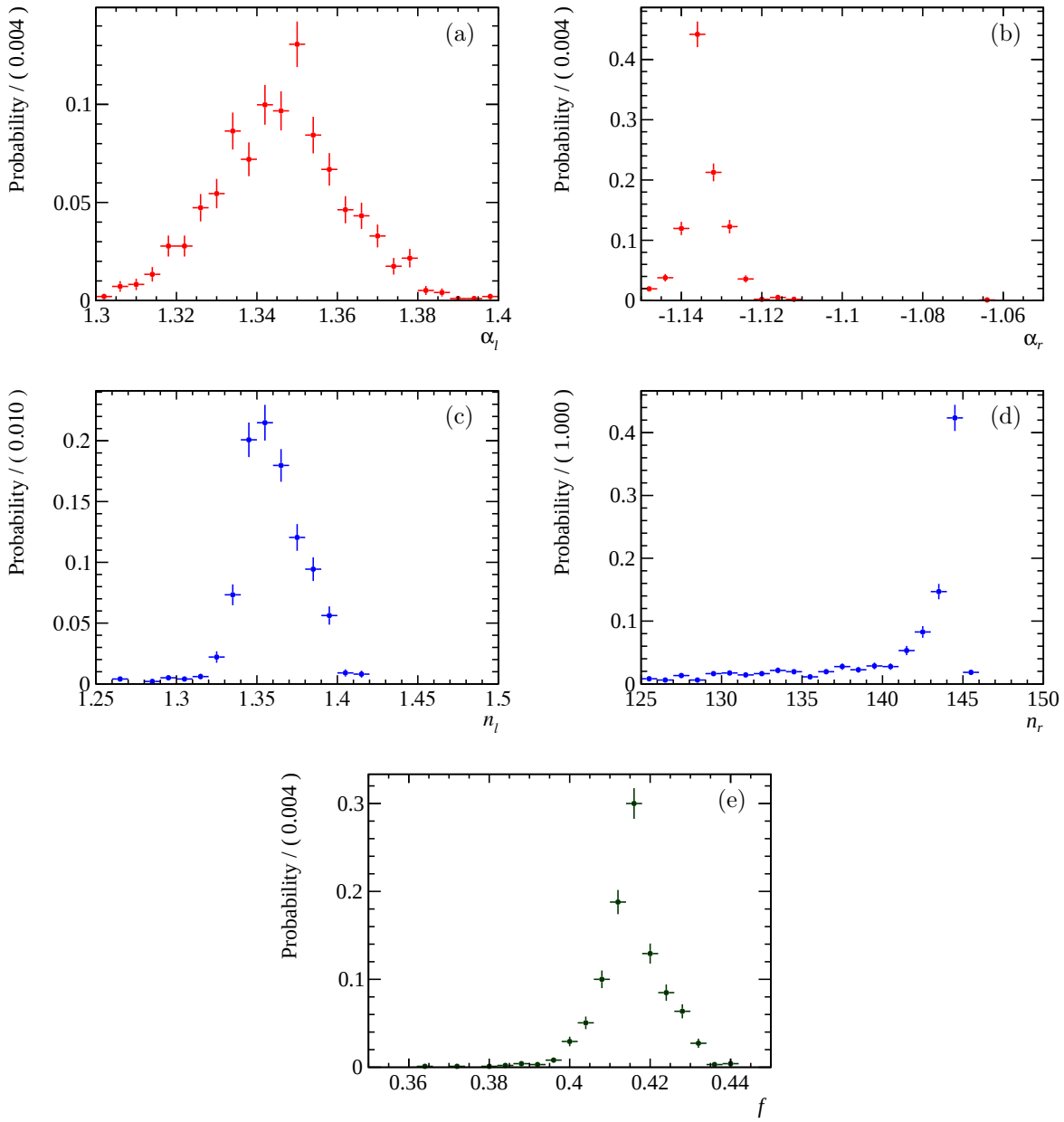


Figure B.2: The distribution of (a) α_l, (b) α_r, (c) n_l, (d) n_r and (e) f for $B^0_s \to \mu^+\mu^-$ in 2011 data. The standard deviation of the distributions is used as systematic uncertainty on the parameters.

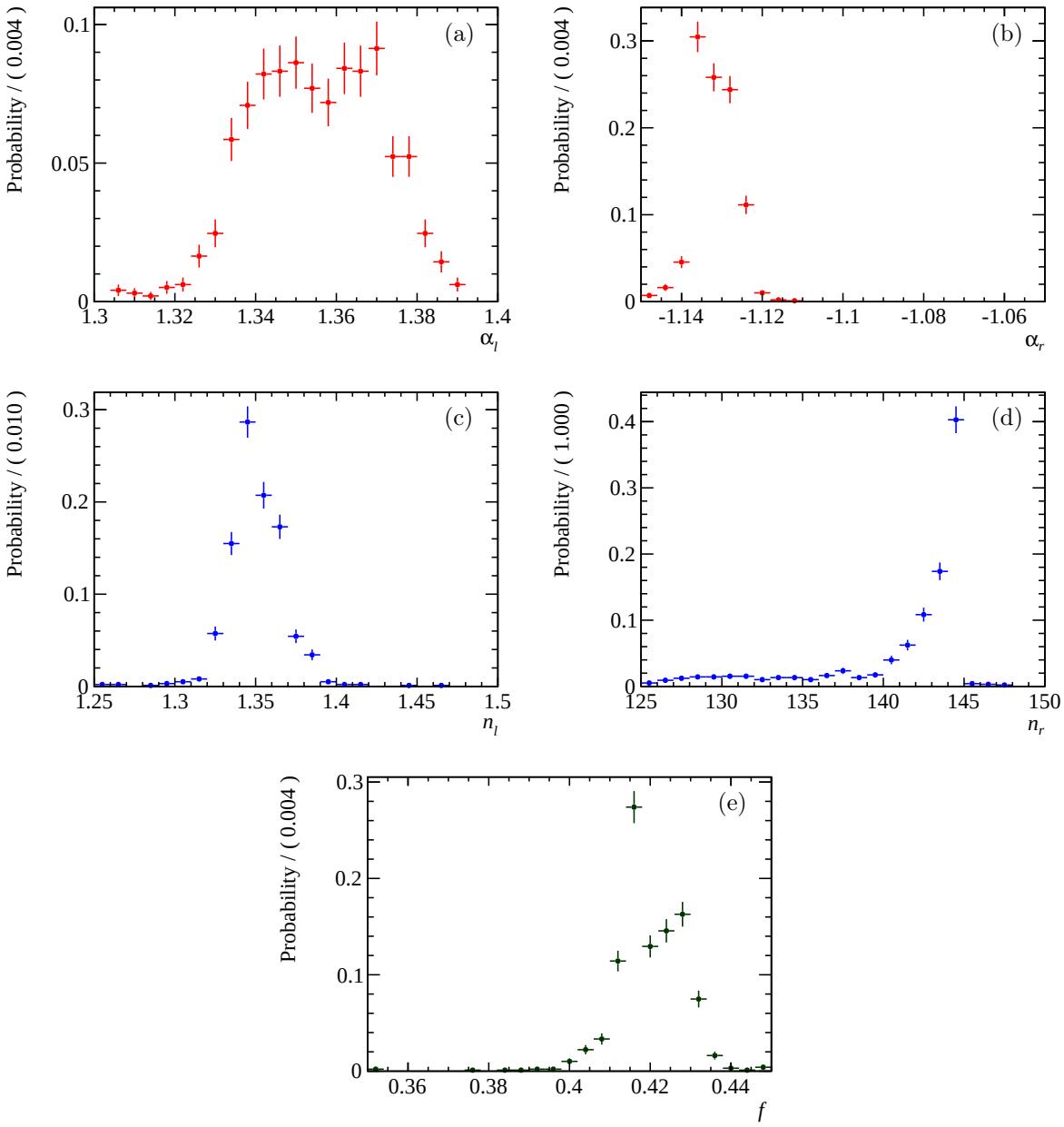


Figure B.3: The distribution of (a) α_l, (b) α_r, (c) n_l, (d) n_r and (e) f for $B^0 \to \mu^+\mu^-$ in 2011 data. The standard deviation of the distributions is used as systematic uncertainty on the parameters.

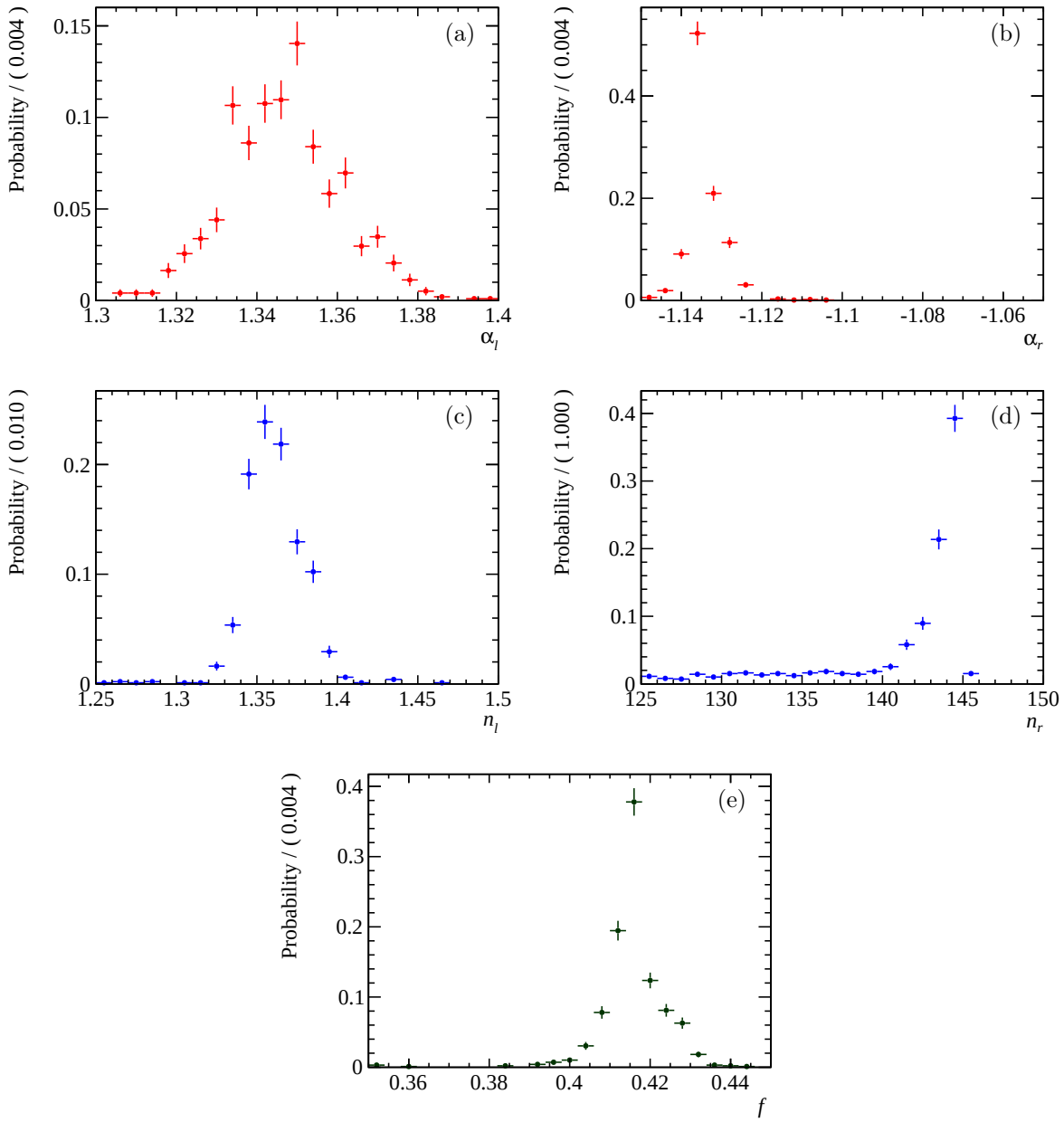


Figure B.4: The distributions of (a) α_l, (b) α_r, (c) n_l, (d) n_r and (e) f for $B^0_s \to \mu^+\mu^-$ in 2012 data. The standard deviation of the distributions is used as systematic uncertainty on the parameters.

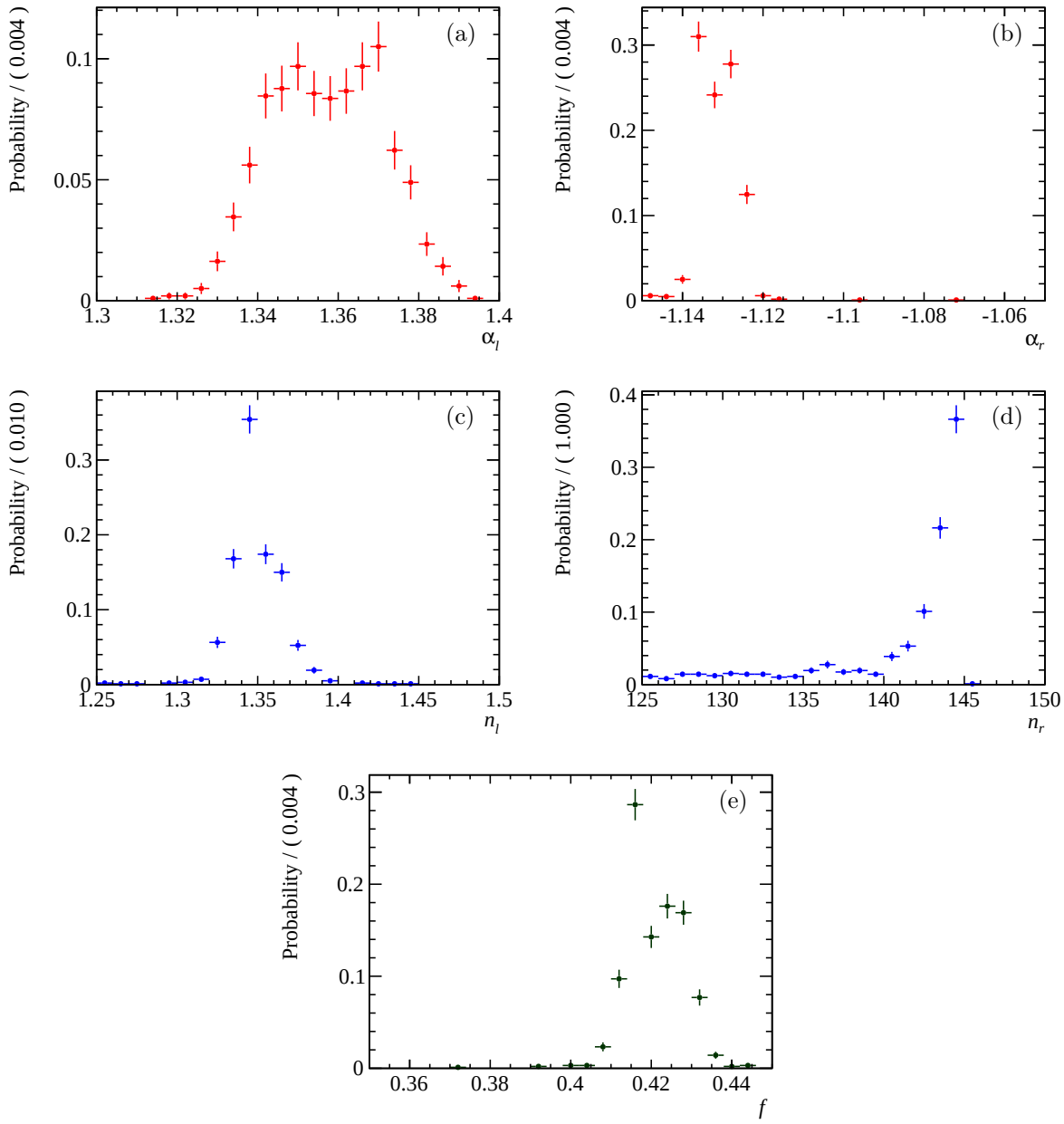


Figure B.5: The distribution of (a) α_l, (b) α_r, (c) n_l, (d) n_r and (e) f for $B^0 \to \mu^+\mu^-$ in 2012 data. The standard deviation of the distributions is used as systematic uncertainty on the parameters.

References

[1] *Convention pour l'établissement d'une Organisation européenne pour la Recherche nucléaire*, CERN, Geneva, 1953.

[2] Photos *Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97-107, arXiv:hep-ph/0506026.

[3] CDF collaboration, *Search for $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ Decays in 3.7 fb^{-1} of $p\overline{p}$ Collisions with CDF II*, CDF Public Note 9892.

[4] ATLAS collaboration, *Measurements of W Boson Yields in Pb+Pb at* 2.76 *TeV/nucleon via single muons with the ATLAS detector*, ATLAS-CONF-2011-078.

[5] ATLAS collaboration, *Measurement of the Z boson Production in Lead-Lead Collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ATLAS Detector*, ATLAS-CONF-2012-052.

[6] CMS and LHCb collaborations, *Combination of results on the rare decays $B^0_{(s)} \to \mu^+\mu^-$ from the CMS and LHCb experiments*, Jul, 2013. CMS-PAS-BPH-13-007, LHCb-CONF-2013-012.

[7] ATLAS collaboration, *Limit on $B_s^0 \to \mu^+\mu^-$ branching fraction based on* 4.9 fb^{-1}*of integrated luminosity*, ATLAS-CONF-2013-076.

[8] LHCb Collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, LHCB-TDR-015.

[9] LHCb collaboration, *Precision luminosity measurements at the LHC*, LHCb-PAPER-2014-047, in preparation.

[10] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[11] ATLAS collaboration, G. Aad *et al.*, *Measurement of the centrality dependence of J/ψ yields and observation of Z production in lead-lead collisions with the ATLAS detector at the LHC*, Phys. Lett. **B697** (2011) 294-312, arXiv:1012.5419.

[12] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1-29, arXiv:1207.7214.

[13] ATLAS collaboration, G. Aad *et al.*, *Search for the decay $B_s^0 \to \mu^+\mu^-$ with the ATLAS detector*, Phys. Lett. **B713** (2012) 387-407, arXiv:1204.0735.

[14] ATLAS collaboration, G. Aad *et al.*, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013) 120-144, arXiv:1307.1432.

[15] ATLAS collaboration, G. Aad *et al.*, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys. Lett. **B726** (2013) 88-119, arXiv:1307.1427.

[16] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, arXiv:1406.3827.

[17] ATLAS Collaboration, G. Aad *et al.*, *Operation and performance of the ATLAS semiconductor tracker*, arXiv:1404.7473.

[18] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

[19] LHCb collaboration, R. Aaij *et al.*, *Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$*, Phys. Lett. **B699** (2011) 330, arXiv:1103.2465.

[20] LHCb collaboration, R. Aaij *et al.*, *Inclusive W and Z production in the forward region at $\sqrt{s} = 7$ TeV*, JHEP **06** (2012) 058, arXiv:1204.1620.

[21] LHCb collaboration, R. Aaij *et al.*, *Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$*, Phys. Lett. **B708** (2012) 55, arXiv:1112.1600.

[22] LHCb collaboration, R. Aaij *et al.*, *Strong constraints on the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$*, Phys. Rev. Lett. **108** (2012) 231801, arXiv:1203.4493.

[23] LHCb collaboration, R. Aaij *et al.*, *A study of the Z production cross-section in pp collisions at $\sqrt{s} = 7$ TeV using tau final states*, JHEP **01** (2013) 111, arXiv:1210.6289.

[24] LHCb collaboration, R. Aaij *et al.*, *First evidence for the decay $B_s^0 \to \mu^+\mu^-$*, Phys. Rev. Lett. **110** (2013) 021801, arXiv:1211.2674.

[25] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, arXiv:1307.5024.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[27] LHCb collaboration, R. Aaij *et al.*, *Observation of $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D_s^{*+}$ decays*, Phys. Rev. **D87** (2013) 112012, arXiv:1304.4530.

[28] LHCb collaboration, R. Aaij *et al.*, *Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV*, Nucl. Phys. **B871** (2013) 1, arXiv:1302.2864.

[29] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[30] LHCb collaboration, R. Aaij *et al.*, *Measurement of the forward W boson production cross-section in pp collisions at* $\sqrt{s} = 7$ *TeV*, arXiv:1408.4354, submitted to JHEP.

[31] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, arXiv:1408.1251, submitted to JINST.

[32] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the* $Z(4430)^-$ *state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.

[33] R. Aaij *et al.*, *Performance of the LHCb calorimeters*, LHCb-DP-2013-004, in preparation.

[34] LHCb collaboration, R. Aaij *et al.*, *Study of forward Z+jet production in pp collisions at* $\sqrt{s} = 7$ *TeV*, JHEP **01** (2014) 033, arXiv:1310.8197.

[35] LHCb collaboration, R. Aaij *et al.*, *Study of* J/ψ *production and cold nuclear matter effects in pPb collisions at* $\sqrt{s_{NN}} = 5$ *TeV*, JHEP **02** (2014) 072, arXiv:1308.6729.

[36] R. Aaij and J. Albrecht, *Muon triggers in the High Level Trigger of LHCb*, LHCb-PUB-2011-017. CERN-LHCb-PUB-2011-017, CERN, Geneva, Sep, 2011.

[37] CDF collaboration, T. Aaltonen *et al.*, *Search for* $B_s \to \mu^+\mu^-$ *and* $B_d \to \mu^+\mu^-$ *Decays with CDF II*, Phys. Rev. Lett. **07** (2011) 191801, arXiv:1107.2304.

[38] CDF collaboration, T. Aaltonen *et al.*, *Search for* $B_s \to \mu^+\mu^-$ *and* $B_d \to \mu^+\mu^-$ *decays with the full CDF Run II data set*, Phys. Rev. **D87** (2013) 072003, arXiv:1301.7048.

[39] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[40] D0 collaboration, V. M. Abazov *et al.*, *Search for the rare decay* $B_s^0 \to \mu^+\mu^-$, Phys. Lett. **B693** (2010) 539-544, arXiv:1006.3469.

[41] D0 collaboration, V. M. Abazov *et al.*, *Search for the rare decay* $B_s \to \mu\mu$, Phys. Rev. **D87** (2013) 072006, arXiv:1301.4507.

[42] L3 collaboration, M. Acciarri *et al.*, *Search for neutral B meson decays to two charged leptons*, Phys. Lett. **B391** (1997) 474-480.

[43] LHCb collaboration, B. Adeva *et al.*, *Road map for selected key measurements from LHCb*, LHCb-PUB-2009-029. CERN-LHCb-PUB-2009-029.

[44] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[45] LHCf collaboration, O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.

[46] A. Affolder *et al.*, *Radiation damage in the LHCb vertex locator*, JINST **8** (2013) P08002, arXiv:1302.5259.

[47] M. Agari *et al.*, *Beetle: A radiation hard readout chip for the LHCb experiment*, Nucl. Instrum. Meth. **A518** (2004) 468-469.

[48] S. Agostinelli *et al.*, GEANT*4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250-303.

[49] UA1 collaboration, C. Albajar *et al.*, *A search for rare B meson decays at the CERN $Sp\overline{p}S$ collider*, Phys. Lett. **B262** (1991) 163-170.

[50] ARGUS collaboration, H. Albrecht *et al.*, *B Meson Decays Into Charmonium States*, Phys. Lett. **B199** (1987) 451.

[51] T. Aliev, N. Pak, and M. Savci, *Rare radiative $B \to \tau^+\tau^-\gamma$ decay*, Phys. Lett. **B424** (1998) 175-184, arXiv:hep-ph/9710304.

[52] J. Allison *et al.*, GEANT*4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[53] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298.

[54] W. Altmannshofer and D. M. Straub, *Cornering New Physics in $b \to s$ Transitions*, JHEP **08** (2012) 121, arXiv:1206.0273.

[55] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[56] LHCb collaboration, S. Amato *et al.*, *LHCb Calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[57] LHCb collaboration, S. Amato *et al.*, *LHCb Magnet: Technical Design Report*, CERN-LHCC-2000-007.

[58] LHCb collaboration, S. Amato *et al.*, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037.

[59] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of early 2012*, arXiv:1207.1158, updated results and plots available at http://www.slac.stanford.edu/xorg/hfag/.

[60] J. Amoraal *et al.*, *Application of vertex and mass constraints in track-based alignment*, Nucl. Instrum. Meth. **A712** (2013) 48-55, arXiv:1207.4756.

[61] TOTEM collaboration, G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[62] LHCb collaboration, R. Antunes Nobrega *et al.*, *LHCb Trigger System: Technical Design Report*, CERN-LHCC-2003-031.

[63] A. Arbey *et al.*, *Implications of a 125 GeV Higgs for supersymmetric models*, Phys. Lett. **B708** (2012) 162-169, arXiv:1112.3028.

[64] A. Arbey, M. Battaglia, F. Mahmoudi, and D. Martinez Santos, *Supersymmetry confronts $B_s \to \mu^+\mu^-$: Present and future status*, Phys. Rev. **D87** (2013) 035026, arXiv:1212.4887.

[65] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[66] S. Arrhenius, *Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte*, Z. Physik. Chem. **4** (1889) 96-116.

[67] BaBar collaboration, B. Aubert *et al.*, *Search for decays of B^0 mesons into e^+e^-, $\mu^+\mu^-$, and $e^\pm\mu^\mp$ final states*, Phys. Rev. **D77** (2008) 032007, arXiv:0712.1516.

[68] European Muon collaboration, J. Aubert *et al.*, *The ratio of the nucleon structure functions F_2^N for iron and deuterium*, Phys. Lett. **B123** (1983) 275.

[69] C. Balazs and C. Yuan, *Soft gluon effects on lepton pairs at hadron colliders*, Phys. Rev. **D56** (1997) 5558-5583, arXiv:hep-ph/9704258.

[70] R. D. Ball *et al.*, *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244-289, arXiv:1207.1303.

[71] R. D. Ball *et al.*, *Parton Distribution Benchmarking with LHC Data*, JHEP **1304** (2013) 125, arXiv:1211.5142.

[72] R. D. Ball *et al.*, *A first unbiased global NLO determination of parton distributions and their uncertainties*, Nucl. Phys. **B838** (2010) 136-206, arXiv:1002.4407.

[73] LHCb collaboration, A. Barbosa Franca *et al.*, *LHCb Inner Tracker: Technical Design Report*, CERN-LHCC-2002-029.

[74] LHCb collaboration, P. Barbosa Marinho *et al.*, *LHCb Muon System: Technical Design Report*, CERN-LHCC-2001-010.

[75] P. Barbosa Marinho *et al.*, *LHCb Outer Tracker: Technical Design Report*, CERN-LHCC-2001-024.

[76] LHCb collaboration, P. Barbosa Marinho *et al.*, *LHCb Vertex locator: Technical Design Report*, CERN-LHCC-2001-011.

[77] G. Battistoni *et al.*, *The* FLUKA *code: Description and benchmarking*, AIP Conf. Proc. **896** (2007) 31-49.

[78] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.

[79] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \to V l^+ l^-$, $V\gamma$ decays*, Nucl. Phys. **B612** (2001) 25-58, arXiv:hep-ph/0106067.

[80] CLEO collaboration, T. Bergfeld *et al.*, *Search for decays of B^0 mesons into pairs of leptons: $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$ and $B^0 \to e^\pm\mu^\mp$*, Phys. Rev. **D62** (2000) 091102, arXiv:hep-ex/0007042.

[81] Particle Data Group, J. Beringer *et al.*, *Review of Particle Physics with 2013 update*, Phys. Rev. **D86** (2013) 010001.

[82] A. Berkien *et al.*, *The front end electronics of the LHCb straw tube tracker*, Nucl. Instrum. Meth. **A617** (2010) 329-331.

[83] C. Bobeth, T. Ewerth, F. Kruger, and J. Urban, *Analysis of neutral Higgs boson contributions to the decays* $\overline{B}_{(s)} \to \ell^+\ell^-$ *and* $\overline{B} \to K\ell^+\ell^-$, Phys. Rev. **D64** (2001) 074014, arXiv:hep-ph/0104284.

[84] C. Bobeth *et al.*, $B_{s,d} \to \ell^+\ell^-$ *in the Standard Model with Reduced Theoretical Uncertainty*, Phys. Rev. Lett. **112** (2014) 101801, arXiv:1311.0903.

[85] C. Bobeth, M. Gorbahn, and E. Stamou, *Electroweak Corrections to* $B_{s,d} \to \ell^+\ell^-$, Phys. Rev. **D89** (2014) 034023, arXiv:1311.1348.

[86] S. Borghi *et al.*, *First spatial alignment of the LHCb VELO and analysis of beam absorber collision data*, Nucl. Instrum. Meth. **A618** (2010) 108-120.

[87] E. Bos and E. Rodrigues, *The LHCb Track Extrapolator Tools*, LHCb-2007-140, CERN-LHCb-2007-140, CERN, Geneva, Nov, 2007.

[88] S. N. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Zeitschrift für Physik **26** (1924), no. 1 178-181.

[89] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[90] CTEQ collaboration, R. Brock *et al.*, *Handbook of perturbative QCD: Version 1.0*, Rev. Mod. Phys. **67** (1995) 157-248.

[91] S. J. Brodsky and H. J. Lu, *Shadowing and Antishadowing of Nuclear Structure Functions*, Phys. Rev. Lett. **64** (1990) 1342.

[92] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125-1144, arXiv:hep-ph/9512380.

[93] O. Buchmueller *et al.*, *The CMSSM and NUHM1 after LHC Run 1*, arXiv:1312.5250.

[94] O. Buchmueller *et al.*, *Implications of Improved Higgs Mass Calculations for Supersymmetric Models*, Eur. Phys. J. **C74** (2014) 2809, arXiv:1312.5233.

[95] A. J. Buras, P. H. Chankowski, J. Rosiek, and L. Slawianowska, $\Delta M_{d,s}$, ${B^0}_{d,s} \to \mu^+\mu^-$ *and* $B \to X_s\gamma$ *in supersymmetry at large* $\tan\beta$, Nucl. Phys. **B659** (2003) 3, arXiv:hep-ph/0210145.

[96] A. J. Buras, R. Fleischer, J. Girrbach, and R. Knegjens, *Probing New Physics with the* $B_s \to \mu^+\mu^-$ *Time-Dependent Rate*, JHEP **1307** (2013) 77, arXiv:1303.3820.

[97] A. J. Buras, J. Girrbach, D. Guadagnoli, and G. Isidori, *On the Standard Model prediction for* $BR(B_{s,d} \to \mu^+\mu^-)$, Eur. Phys. J. **C72** (2012) 2172, arXiv:1208.0934.

[98] S. Burby and C. Maxwell, *Direct extraction of QCD $\Lambda(\overline{MS})$ from e^+e^- jet observables*, Nucl. Phys. **B609** (2001) 193-224, arXiv:hep-ph/0011203.

[99] H. Burkhardt and P. Grafstrom, *Absolute luminosity from machine parameters*, CERN-LHC-PROJECT-REPORT-1019, 2007.

[100] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531-533.

[101] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189.

[102] O. Callot, *Downstream Pattern Recognition*, LHCb-2007-026, CERN-LHCb-2007-026.

[103] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, LHCb-2007-015, CERN-LHCb-2007-015.

[104] J. Campbell, J. Huston, and W. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Rept. Prog. Phys. **70** (2007), no. 1 89.

[105] A. Carbone *et al.*, *Invariant mass line shape of $B \to PP$ decays at LHCb*, LHCb-PUB-2009-031. CERN-LHCb-PUB-2009-031.

[106] J. Casas and C. Munoz, *A Natural Solution to the μ Problem*, Phys. Lett. **B306** (1993) 288-294, arXiv:hep-ph/9302227.

[107] BELLE collaboration, M. Chang *et al.*, *Search for $B^0 \to \ell^+\ell^-$ at BELLE*, Phys. Rev. **D68** (2003) 111101, arXiv:hep-ex/0309069.

[108] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[109] CMS collaboration, S. Chatrchyan *et al.*, *Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays in pp collisions at 7 TeV*, Phys. Rev. Lett. **107** (2011) 191802, arXiv:1107.5834.

[110] CMS collaboration, S. Chatrchyan *et al.*, *Study of Z boson production in PbPb collisions at nucleon-nucleon centre of mass energy = 2.76 TeV*, Phys. Rev. Lett. **106** (2011) 212301, arXiv:1102.5435.

[111] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30-61, arXiv:1207.7235.

[112] CMS collaboration, S. Chatrchyan *et al.*, *Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays*, JHEP **04** (2012) 033, arXiv:1203.3976.

[113] CMS collaboration, S. Chatrchyan *et al.*, *Study of W boson production in PbPb and pp collisions at $\sqrt{s_{NN}} = 2.76$ TeV*, Phys. Lett. **B715** (2012) 66-87, arXiv:1205.6334.

[114] CMS collaboration, S. Chatrchyan *et al.*, *Measurement of the $B_s \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ with the CMS Experiment*, Phys. Rev. Lett. **111** (2013) 101804, arXiv:1307.5025.

[115] CMS collaboration, S. Chatrchyan *et al.*, *Search for a Higgs boson decaying into a Z and a photon in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, Phys. Lett. **B726** (2013) 587-609, arXiv:1307.5515.

[116] CMS collaboration, S. Chatrchyan *et al.*, *Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs*, Phys. Rev. Lett. **110** (2013) 081803, arXiv:1212.6639.

[117] K. G. Chetyrkin, M. Misiak, and M. Munz, *Weak radiative B meson decay beyond leading logarithms*, Phys. Lett. **B400** (1997) 206-219, arXiv:hep-ph/9612313.

[118] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[119] J. C. Collins and D. E. Soper, *Back-To-Back Jets in QCD*, Nucl. Phys. **B193** (1981) 381.

[120] J. C. Collins and D. E. Soper, *Back-To-Back Jets: Fourier Transform from B to K-Transverse*, Nucl. Phys. **B197** (1982) 446.

[121] J. C. Collins, D. E. Soper, and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, Nucl. Phys. **B250** (1985) 199.

[122] B. Combridge, *Associated production of heavy flavour states in pp and $p\overline{p}$ interactions: Some QCD estimates*, Nucl. Phys. **B151** (1979) 429-456.

[123] G. Corti *et al.*, *Software for the LHCb experiment*, IEEE Trans. Nucl. Sci. **53** (2006) 1323-1328.

[124] N. Craig and S. Thomas, *Exclusive Signals of an Extended Higgs Sector*, JHEP **11** (2012) 083, arXiv:1207.4835.

[125] K. De Bruyn *et al.*, *Branching Ratio Measurements of B_s Decays*, Phys. Rev. **D86** (2012) 014027, arXiv:1204.1735.

[126] M. De Cian, U. Straumann, O. Steinkamp, and N. Serra, *Track Reconstruction Efficiency and Analysis of $B^0 \to K^{*0}\mu^+\mu^-$ at the LHCb Experiment*, PhD thesis, University of Zurich, Sep, 2013.

[127] D. de Florian, R. Sassot, P. Zurita, and M. Stratmann, *Global analysis of nuclear parton distributions*, Phys. Rev. **D85** (2012) 074028, arXiv:1112.6324.

[128] LHCb collaboration, H. Dijkstra, H. J. Hilke, T. Nakada, and T. Ypsilantis, *LHCb Letter of Intent*, LHCb-95-001.

[129] P. A. M. Dirac, *On the theory of quantum mechanics*, Proceedings of the Royal Society of London. Series A **112** (1926), no. 762 661-677.

[130] G. Dissertori, I. Knowles, and M. Schmelling, *Quantum Chromodynamics: High Energy Experiments and Theory*, OUP Oxford, 2009.

[131] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641-653.

[132] G. Dvali, G. Giudice, and A. Pomarol, *The μ-problem in theories with gauge mediated supersymmetry breaking*, Nucl. Phys. **B478** (1996) 31-45, arXiv:hep-ph/9603238.

[133] J. R. Ellis *et al.*, *Higgs Bosons in a Nonminimal Supersymmetric Model*, Phys. Rev. **D39** (1989) 844.

[134] J. R. Ellis, K. A. Olive, and V. C. Spanos, *On the interpretation of $B_s \to \mu^+\mu^-$ in the CMSSM*, Phys. Lett. **B624** (2005) 47-59, arXiv:hep-ph/0504196.

[135] R. Ellis, W. Stirling, and B. Webber, *QCD and Collider Physics*, Cambridge University Press, 2003.

[136] C. Elsasser, *Extension of the TT Test Stand with a Pulsed Focused Infrared Laser*, Bachelor thesis, University of Zurich, Nov, 2009.

[137] C. Elsasser, *Determination of the Invariant Mass Distribution for $B^0_{(s)} \to \mu^+\mu^-$ at the LHCb Experiment*, Master thesis, University of Zurich, Aug, 2011.

[138] C. Elsasser, *The LHCb Silicon Tracker*, JINST **9** (2014), no. 01 C01009.

[139] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321-323.

[140] K. Eskola, *Global analysis of nuclear PDFs - latest developments*, Nucl. Phys. **A 910-911** (2013) 163, arXiv:1209.1546.

[141] K. Eskola, H. Paukkunen, and C. Salgado, *EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions*, JHEP **04** (2009) 065, arXiv:0902.4154.

[142] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[143] H. Feick, *Radiation tolerance of silicon particle detectors for high-energy physics experiments*, PhD thesis, University of Hamburg, Sep, 1997, DESY-F35D-97-08.

[144] E. Fermi, *Zur Quantelung des idealen einatomigen Gases*, Zeitschrift für Physik **36** (1926), no. 11-12 902-912.

[145] E. Fermi, *An attempt of a theory of beta radiation*, Z. Phys. **88** (1934) 161-177.

[146] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, FLUKA*: A multi-particle transport code (Program version 2005)*, CERN-2005-010, SLAC-R-773, INFN-TC-05-11.

[147] R. Fleischer, N. Serra, and N. Tuning, *Tests of Factorization and SU(3) Relations in B Decays into Heavy-Light Final States*, Phys. Rev. **D83** (2011) 014017, arXiv:1012.2784.

[148] R. W. Forty, *Track Seeding*, LHCb-2001-109.

[149] L. Frankfurt, V. Guzey, and M. Strikman, *Leading twist nuclear shadowing phenomena in hard processes with nuclei*, Phys. Rept. **512** (2012) 255-393, arXiv:1106.2091.

[150] RD2 Collaboration, E. Fretwurst *et al.*, *Reverse annealing of the effective impurity concentration and long-term operational scenario for silicon detectors in future collider experiments*, Nucl. Instrum. Methods Phys. Res. , A **342** (1993) 119-125. 15 p.

[151] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, FEWZ *2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput. Phys. Commun. **182** (2011) 2388-2403, arXiv:1011.3540.

[152] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Il Nuovo Cimento **4** (1956) 848-866.

[153] G. Giudice and A. Masiero, *A Natural Solution to the μ Problem in Supergravity Theories*, Phys. Lett. **B206** (1988) 480-484.

[154] S. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 - 588.

[155] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285-1292.

[156] R. Glauber, *Cross-sections in deuterium at high-energies*, Phys. Rev. **100** (1955) 242-248.

[157] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003, CERN-LHCb-PUB-2011-003, LHCb-INT-2010-053.

[158] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016, CERN-LHCb-PUB-2011-016, LHCb-INT-2011-030, LHCb-INT-2011-030.

[159] R. Gluckstern, *Uncertainties in track momentum and direction, due to multiple scattering and measurement errors*, Nucl. Instrum. Meth. **24** (1963) 381-389.

[160] E. Golowich *et al.*, *Relating B_s Mixing and $B_s \to \mu^+\mu^-$ with New Physics*, Phys. Rev. **D83** (2011) 114017, arXiv:1102.0009.

[161] H. Gong, Y.-D. Yang, and X.-B. Yuan, *Constraints on anomalous tcZ coupling from $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s \to \mu^+\mu^-$ decays*, JHEP **1305** (2013) 062, arXiv:1301.7535.

[162] V. Gribov and L. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438-450.

[163] C. Grojean, G. Servant, and J. D. Wells, *First-order electroweak phase transition in the standard model with a low cutoff*, Phys. Rev. **D71** (2005) 036001, arXiv:hep-ph/0407019.

[164] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585-587.

[165] V. Guzey *et al.*, *Massive neutral gauge boson production as a probe of nuclear modifications of parton distributions at the LHC*, Eur. Phys. J. **A49** (2013) 35, arXiv:1212.5344.

[166] G. Haefeli *et al.*, *The LHCb DAQ interface board TELL1*, Nucl. Instrum. Meth. **A560** (2006) 494-502.

[167] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, 1984.

[168] T. Hermann, M. Misiak, and M. Steinhauser, *Three-loop QCD corrections to $B_s \to \mu^+\mu^-$*, JHEP **12** (2013) 097, arXiv:1311.1347.

[169] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508-509.

[170] M. Hirai, S. Kumano, and T.-H. Nagai, *Determination of nuclear parton distribution functions and their uncertainties in next-to-leading order*, Phys. Rev. **C76** (2007) 065207, arXiv:0709.3038.

[171] W.-S. Hou, M. Kohda, and F. Xu, *Implication of possible observation of enhanced $B_d^0 \to \mu^+\mu^-$ decay*, Phys. Rev. **D87** (2013) 094005, arXiv:1302.1471.

[172] W. D. Hulsbergen, *The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, Nucl. Instrum. Meth. **A600** (2009), no. 2 471 - 477.

[173] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566-575, arXiv:physics/0503191.

[174] R. Jacobsson, *Performance of the LHCb Detector During the LHC Proton Runs 2010 - 2012*, Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE (2013) 1479 - 1486.

[175] U. Kerzel, *The LHCb RICH detectors*, J. Phys. Conf. Ser. **110** (2008) 092014.

[176] CMS Collaboration, V. Khachatryan *et al.*, *Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs*, arXiv:1405.3455.

[177] CMS Collaboration, V. Khachatryan *et al.*, *Observation of the diphoton decay of the Higgs boson and measurement of its properties*, arXiv:1407.0558.

[178] D. Kharzeev and H. Satz, *The Quantum mechanical origin of nuclear shadowing*, Phys. Lett. **B327** (1994) 361-368.

[179] K. Kleinknecht, *Detectors for particle radiation*, Cambridge University Press, 2nd ed., 1998.

[180] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652-657.

[181] K. Kovarik *et al.*, *CTEQ nuclear parton distribution functions*, PoS **DIS2013** (2013) 274, arXiv:1307.3454.

[182] K. Kovarik *et al.*, *Nuclear corrections in neutrino-nucleus DIS and their compatibility with global NPDF analyses*, Phys. Rev. Lett. **106** (2011) 122301, arXiv:1012.0286.

[183] H.-L. Lai *et al.*, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241.

[184] J. Laiho, E. Lunghi, and R. S. Van de Water, *Lattice QCD inputs to the CKM unitarity triangle analysis*, Phys. Rev. **D81** (2010) 034503, arXiv:0910.2928.

[185] F. Landry, R. Brock, P. M. Nadolsky, and C. Yuan, *Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism*, Phys. Rev. **D67** (2003) 073016, arXiv:hep-ph/0212159.

[186] G. Lanfranchi *et al.*, *The Muon Identification Procedure of the LHCb Experiment for the First Data*, LHCb-PUB-2009-013. CERN-LHCb-PUB-2009-013, CERN, Geneva, Aug, 2009.

[187] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152-155.

[188] LHCb collaboration, *Inclusive low mass Drell-Yan production in the forward region at* $\sqrt{s} = 7\,\text{TeV}$, LHCb-CONF-2012-013.

[189] LHCb collaboration, *Graphical comparison of the LHCb measurements of W and Z boson production with ATLAS and CMS*, LHCb-CONF-2013-005.

[190] LHCb collaboration, *Measurement of the cross section for* $Z \to \mu^+\mu^-$ *production with* $1.0\,fb^{-1}$ *of pp collisions at* $\sqrt{s} = 7\,\text{TeV}$, LHCb-CONF-2013-007.

[191] Y. Li and F. Petriello, *Combining QCD and electroweak corrections to dilepton production in the framework of the* FEWZ *simulation code*, Phys. Rev. **D86** (2012) 094034, arXiv:1208.5967.

[192] W. Liu, C.-X. Yue, and H.-D. Yang, *Rare decays* $B_s \to \ell^+\ell^-$ *and* $B \to K\ell^+\ell^-$ *in the topcolor-assisted technicolor model*, Phys. Rev. **D79** (2009) 034008, arXiv:0901.3463.

[193] H. E. Logan and U. Nierste, $B_{s,d} \to \ell^+\ell^-$ *in a two Higgs doublet model*, Nucl. Phys. **B586** (2000) 39-55, arXiv:hep-ph/0004139.

[194] F. Mahmoudi, *Indirect search for New Physics: complementarity to direct searches*, arXiv:1310.2556.

[195] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189-285, arXiv:0901.0002.

[196] CERN-ROSE/RD48 collaboration, M. Moll, E. Fretwurst, and G. Lindstrom, *Leakage current of hadron irradiated silicon detectors - material dependence*, Nucl. Instrum. Meth. **A426** (1999) 87-93.

[197] M. Moll, E. Fretwurst, G. Lindstrom, and M. Kuhnke, *Relation between microscopic defects and macroscopic changes in silicon detector properties after hadron irradiation*, Nucl. Instrum. Meth. **B186** (2002) 100-110.

[198] M. Moll, *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties*, PhD thesis, University of Hamburg, Dec, 1999, DESY-THESIS-1999-040.

[199] G. P. Mueller and C. S. Guenzer, *Simulation of cascade damage in silicon*, IEEE Trans. Nucl. Sci. **27** (1980), no. 6 1474-1477.

[200] P. M. Nadolsky *et al.*, *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, arXiv:0802.0007.

[201] P. Nason, S. Dawson, and R. K. Ellis, *The One Particle Inclusive Differential Cross-Section for Heavy Quark Production in Hadronic Collisions*, Nucl. Phys. **B327** (1989) 49-92.

[202] M. Needham and J. Van Tilburg, *Performance of the track matching*, LHCb-2007-020, CERN-LHCb-2007-020.

[203] K. Nishijima, *Charge independence theory of v particles*, Progress of Theoretical Physics **13** (1955), no. 3 285-304.

[204] E. Noether, *Invarianten beliebiger Differentialausdrücke*, Gött. Nachr. (1918) 37-44.

[205] E. Norrbin and T. Sjöstrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000), no. 1 137-161.

[206] H. Paukkunen, *Nuclear PDFs in the beginning of the LHC era*, arXiv:1401.2345.

[207] J. Pinfold, *The MoEDAL Experiment – Searching for Highly Ionizing Particles at the LHC*, Progress in High Energy Physics and Nuclear Safety – NATO Science for Peace and Security Series B: Physics and Biophysics (2009) 217-226.

[208] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356-369, arXiv:physics/0402083.

[209] L. Randall and R. Sundrum, $b \to s\gamma$ *and* $B_s \to \mu^+\mu^-$ *in extended technicolor models*, Phys. Lett. **B312** (1993) 148-154, arXiv:hep-ph/9305289.

[210] A. Read, *Presentation of search results: the* CL_s *technique*, J. Phys. G: Nucl. Part. Phys **28** (2002) 2693.

[211] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577-584, arXiv:physics/0408124.

[212] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367-377.

[213] C. Salgado *et al.*, *Proton-Nucleus Collisions at the LHC: Scientific Opportunities and Requirements*, J. Phys. **G39** (2012) 015010, arXiv:1105.3919.

[214] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

[215] I. Schienbein *et al.*, *PDF Nuclear Corrections for Charged and Neutral Current Processes*, Phys. Rev. **D80** (2009) 094004, arXiv:0907.2357.

[216] T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA *6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[217] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to* PYTHIA *8.1*, Comput. Phys. Commun. **178** (2008) 852-867, arXiv:0710.3820.

[218] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[219] H. Spieler, *Introduction to radiation-resistant semiconductor devices and circuits*, AIP Conf. Proc. **390** (1997) 23-49.

[220] J. Srour, S. Chen, S. Othmer, and R. Hartmann, *Radiation damage coefficients for silicon depletion regions*, IEEE Trans. Nucl. Sci. **26** (1979), no. 6 4783-4791.

[221] M. S. Sze and K. K. Ng, *Physics of Semiconductor Devices*, Wiley, 3rd ed., 2006.

[222] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. **B61** (1973) 455-468.

[223] M. Tobin, *Performance of the LHCb Tracking Detectors*, PoS **Vertex2012** (2013) 047.

[224] N. Tuning, *Detailed performance of the Outer Tracker at LHCb*, JINST **9** (2014), no. 01 C01040.

[225] N. Tuning *et al.*, *Ageing in the LHCb outer tracker: Aromatic hydrocarbons and wire cleaning*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **656** (2011), no. 1 45-50.

[226] S. van der Meer, *Calibration of the Effective Beam Height in the ISR*, CERN-ISR-PO-68-31, 1968.

[227] B. Van Zeghbroeck, *Principles of Semiconductor Devices*, 2006.

[228] S. Viret, C. Parkes, and M. Gersabeck, *Alignment procedure of the LHCb Vertex Detector*, Nucl. Instrum. Meth. **A596** (2008) 157-163, `arXiv:0807.5067`.

[229] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264-1266.

[230] S. Weinberg, *New approach to the renormalization group*, Phys. Rev. **D8** (1973) 3497-3509.

[231] L. Weinstein *et al.*, *Short Range Correlations and the EMC Effect*, Phys. Rev. Lett. **106** (2011) 052301, `arXiv:1009.5666`.

[232] J. Wenninger, *Energy Calibration of the LHC Beams at 4 TeV*, CERN-ATS-2013-040, Geneva, May, 2013.

[233] M. Williams *et al.*, *The HLT2 Topological Lines*, LHCb-PUB-2011-002, CERN-LHCb-PUB-2011-002.

[234] R. Wunstorf, *Systematische Untersuchungen zur Strahlenresistenz von Silizium-Detektoren für die Verwendung in Hochenergiephysik-Experimenten*, PhD thesis, University of Hamburg, Nov, 1992.

[235] K. Wyllie *et al.*, *Silicon detectors and electronics for pixel hybrid photon detectors*, Nucl. Instrum. Meth. **A530** (2004) 82-86.

[236] Y. Xie, *Short track reconstruction with VELO and TT*, LHCb-2003-100.

[237] *The History of CERN*, http://timeline.web.cern.ch/timelines/The-history-of-CERN [Online, accessed November 22, 2013].

[238] *LHCb homepage*, http://lhcb.cern.ch [Online; accessed November, 14, 2013].

[239] *LHCb operations plots*, http://lhcb-operationsplots.web.cern.ch/ [Online; accessed February, 23, 2014].

[240] *Properties of CMS-OB2-SSSD sensor type*, http://lhcb.physik.uzh.ch/tt/sensorprob/S9153-04B.pdf [Online; accessed March, 6, 2014].

[241] *Silicon Tracker Homepage*, http://lhcb.physik.uzh.ch/ST/ [Online; accessed October 20, 2013].

List of Figures

List of Tables